STATISTICS

AND

PHYSICAL OCEANOGRAPHY

National Research Council

STATISTICS

AND

PHYSICAL OCEANOGRAPHY

Panel on Statistics and Oceanography
Committee on Applied and Theoretical Statistics
Board on Mathematical Sciences
Commission on Physical Sciences, Mathematics, and Applications
National Research Council

National Academy Press
Washington, D.C. 1993

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Frank Press is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Robert M. White is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Frank Press and Dr. Robert M. White are chairman and vice chairman, respectively, of the National Research Council.

Support for this project was provided by the Office of Naval Research.

Additional copies of this report are available from:
Board on Mathematical Sciences
National Research Council
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Printed in the United States of America

PANEL ON STATISTICS AND OCEANOGRAPHY

DUDLEY B. CHELTON, Oregon State University, *Cochair*
WILLIAM F. EDDY, Carnegie Mellon University, *Cochair*
RICHARD DEVEAUX, Princeton University
RAISA FELDMAN, University of California at Santa Barbara
ROMAN E. GLAZMAN, Jet Propulsion Laboratory, California Institute of Technology
ANNALISA GRIFFA, University of Miami
KATHRYN A. KELLY, Woods Hole Oceanographic Institution
GORDON J. MACDONALD, University of California at San Diego
MURRAY ROSENBLATT, University of California at San Diego
BORIS ROZOVSKII, University of Southern California

Staff
JOHN R. TUCKER, Senior Program Officer

COMMITTEE ON APPLIED AND THEORETICAL STATISTICS

WILLIAM F. EDDY, Carnegie Mellon University, *Chair*
YVONNE BISHOP, U.S. Department of Energy
MARY ELLEN BOCK, Purdue University
MARJORIE G. HAHN, Tufts University
DOUGLAS M. HAWKINS, University of Minnesota
DAVID G. HOEL, Medical University of South Carolina
JON R. KETTENRING, Bellcore
KARL E. PEACE, Biopharmaceutical Research Consultants
STEPHEN M. POLLOCK, University of Michigan
DARYL PREGIBON, AT&T Bell Laboratories

Staff
JOHN R. TUCKER, Senior Program Officer

COMMISSION ON PHYSICAL SCIENCES, MATHEMATICS, AND APPLICATIONS

PREFACE

This report was prepared in response to a request from the Office of Naval Research to the National Research Council's Committee on Applied and Theoretical Statistics. It describes research opportunities in statistics and applied probability arising in physical oceanographic applications. The report is expository, with the intended audience being statisticians and quantitatively literate people with a background in statistical applications to science, as well as federal agency representatives interested in encouraging such cross-disciplinary research.

In producing this report, the panel had to surmount communication and comprehension difficulties to truly understand, e.g., what someone from another discipline had expressed. One result was an appreciation of just how difficult it is to engage in truly collaborative, cross-disciplinary work. Another result was an insight into what strategies will (and will not) be likely to succeed in performing such work. The panel believes understanding and appreciating these matters are as important to the encouragement and accomplishment of statistical research in physical oceanography as are the descriptions of statistical research opportunities discussed in Chapters 2 through 8. Accordingly, Chapter 9 gives the panel's conclusions, observations, and suggestions on encouraging successful collaborations between statisticians and oceanographers.

The panel gratefully acknowledges the support of the Office of Naval Research in this project and expresses appreciation to all of the people who provided information that aided the panel in the preparation of this report. They include Mark Abbott, Andrew Bennett, Hans Graber, Greg Holloway, Ricardo Matano, Robert N. Miller, Leonid Piterbarg, Michael Schlax, P. Ted Strub, V. Zlotnicki, and four anonymous reviewers who offered insightful comments and suggestions. In particular, L. Piterbarg helped write Chapter 3, P. Strub helped write Chapter 4, M. Abbott helped write Chapter 5, R. Miller and V. Zlotnicki helped write Chapter 6, and H. Graber helped write Chapter 7. The panel also gratefully acknowledges the editorial help of John Tucker and Susan Maurizi in preparing the report.

Comments on the report are welcome, as are suggestions for future topics on which similar reports might help to provide useful cross-disciplinary bridges. All such remarks should be directed to John Tucker at the Board on Mathematical Sciences, National Research Council, Washington, D.C.

CONTENTS

1
OVERVIEW

INTRODUCTION

Purpose and Scope of This Report

Research in oceanography has historically been pursued to better understand the oceans as, for example, avenues to exploration, routes for commerce, theaters for military operations, and components in the weather system. Today this research is also done in conjunction with studies on major issues such as global climate, environmental change, and biodiversity, among many others. Statistical techniques have always been important in the analysis of oceanographic data. With the recent introduction of oceanographic observational mechanisms that yield much larger quantities of data than ever before, statistical considerations have gained even more prominence in oceanographic research contexts. Yet disciplinary distinctions have limited interactions across discipline boundaries in many national and global research areas (NRC, 1987, 1990a); traditional statistics and oceanography are not exceptions. To stimulate progress on important research questions now arising at this interface, more cross-disciplinary efforts between statistics and oceanography are needed. This report is thus presented to help encourage successful collaborations between statistics and oceanography that are focused on potentially fruitful cross-disciplinary research areas.

The report was prepared in response to a request from the Mathematical Sciences Division of the Office of Naval Research for a cross-disciplinary report describing basic research questions in statistics and applied probability motivated by oceanographic applications. The request reflects ONR's desire to call such questions to the attention of research statisticians and to develop stronger interactions between the statistics and oceanography research communities. A panel of five oceanographers and five statisticians was convened by the Committee on Applied and Theoretical Statistics of the National Research Council to produce the report. The charge to the panel was to survey crossover areas between statistics and oceanography of greatest potential value (with respect to important oceanographic questions) and to recommend statistical research opportunities. The panel met in April 1992 and again in August 1992. It quickly became apparent that a comprehensive summary of statistical research opportunities addressing all disciplines of oceanography would exceed the project time and budget constraints. This report is therefore limited to a discussion of statistical research opportunities arising in physical oceanography.

Lest the limited scope of this report be misconstrued as a statement of the unimportance of statistical analysis to biological, chemical, and geological oceanography, the panel emphasizes that there are numerous opportunities for statisticians to work in those disciplines as well. For example, recent interest in the carbon cycle has focused attention on the spatial and seasonal distributions of phytoplankton pigment concentration in the ocean. These data, obtained by satellite, exhibit all the challenges of sparsity and

incompleteness shared by the other data sets discussed in this chapter, and furthermore exhibit temporal and spatial correlation. An eventual question to address is the role of phytoplankton distribution in climate change, but first a quantitative analysis of the distribution itself is necessary. Factors such as bathymetry, nutrients, eddy kinetic energy, wind stress, cloud cover, meltwater formation, and Ekman upwelling are believed to be potential influences on the phytoplankton distribution, but the relationships are as yet unknown. Currently available data on many of these factors are sparse, and a great deal of spatial and temporal aggregation is necessary in order to assess such potential relationships. Future satellite observations are expected to ameliorate the data issues basic to the study of these important biological and chemical oceanographic processes, but the statistical problems discussed in Chapters 2 through 8 will remain the same.

In physical oceanography, the development and application of statistical analysis techniques are somewhat more advanced than in other disciplines of oceanography. In large part, a greater need for sophisticated statistical techniques in physical oceanography has been driven by rapid technological advances over the past 30 years or so that have resulted in larger volumes of observational data spanning a broader range of space and time scales than are available in the other oceanographic disciplines. There has also been intensive development of a theoretical foundation to explain the observations. As a result of these two parallel efforts and recognition of the importance of physical oceanographic processes in many of today's important global issues, there are many significant opportunities for applications of statistics, both where descriptive analyses of the observational data are needed and where there is a need to relate observations to theory. Even the limited scope of physical oceanography presents a rather daunting task for those who would explore it, since the discipline encompasses a very broad range of topics. Input to the panel was sought and was generously provided by several outside experts (see the preface) to broaden the span of topics outlined in this report.

It should be emphasized at the outset that statistical analyses of physical oceanographic data have not been developed in total isolation from developments in the field of statistics. On the contrary, statistical techniques are already used to an unusual degree of sophistication compared with their use in some other scientific disciplines, partly because of the need to develop techniques to understand the almost overwhelming quantity of observational data available. In this regard, physical oceanography has benefitted from the parallel development of techniques of statistical analysis in the field of atmospheric sciences, in which researchers also need to interpret the large volumes of atmospheric data available. Physical oceanographers are generally well versed in traditional and many modern statistical analysis techniques. In addition, several books and monographs have been written specifically on applications of statistical techniques in the atmospheric sciences and physical oceanography (e.g., Gandin, 1965; Thiebaux and Pedder, 1987; Preisendorfer, 1988; Daley, 1991; Ghil and Malanotte-Rizzoli, 1991; Bennett, 1992). Many statistical techniques tailored to specific analyses of oceanographic data have also been published in journal articles.

This report consists of a collection of sections (Chapters 2 through 8) outlining research problems that the panel believes could serve as fruitful areas for collaboration between statisticians and oceanographers. In Chapter 9, the panel presents its conclusions, observations, and suggestions on encouraging successful collaborations between statistics and

oceanography. As noted above, physical oceanographic research encompasses a very broad range of topics. Not all of these subdisciplines are represented by the five oceanographers on the panel. This report should therefore be viewed as a compendium of research interests reflecting the viewpoints of the oceanographers on the panel. This somewhat parochial bias should be kept in mind when using this report to identify potential crossover areas between statistics and physical oceanography; there are likely many statistical research opportunities that have not been identified in the report. Notwithstanding these limitations, the panel believes that the report represents a good first step toward encouraging interaction between statisticians and physical oceanographers to the mutual benefit of both disciplines.

Oceanography — A Brief Sketch

The birth of oceanography as a science can be traced back to 1769, when Benjamin Franklin contributed significantly to scientific knowledge of the oceans by charting sea surface temperature in the North Atlantic and noting that the maximum flow of the Gulf Stream (which had been known to exist and had been used for navigation for a long time) occurred where surface temperatures began dropping rapidly for a ship traveling from the New World to the Old World. Further scientific surveys of the ocean were conducted during this same era by Captain James Cook, who set sail from England in 1772 with the primary goal of making a detailed map of the Pacific Ocean and learning the natural history of the Pacific region. Fontaine Maury is generally credited as the founding father of international oceanographic science. As a U.S. Navy officer, Maury published an atlas (Maury, 1855) based on a worldwide compilation of data taken from ship logbooks. The culmination of this era of scientific exploration of the ocean was the historic voyage of the HMS *Challenger* funded in 1873 by Great Britain to collect detailed measurements of the physical, biological, and chemical characteristics of the world oceans. The 4-year expedition resulted in some 50 volumes of reports published between 1890 and 1895.

The 20th century has witnessed a dramatic expansion of oceanographic research. At the beginning of the century, most of the deep ocean was thought to be relatively quiescent. Except for moderate seasonal variability, it was generally believed that the circulation near the surface of the oceans was relatively constant and large scale. Scripps Institution of Oceanography was founded in 1903 and the Woods Hole Oceanographic Institution was established in 1930. As a result of new technological developments, it became possible to measure physical, chemical, and biological characteristics from the sea surface to the ocean bottom. Dedicated research vessels set out to systematically map the three-dimensional physical, chemical, and biological characteristics of the world ocean on a coarse spatial grid. Although tremendous progress was made in the field of oceanography prior to World War II, it was still possible to summarize existing knowledge in all three disciplines (physical, biological, and chemical) in a single book (Sverdrup et al., 1942).

The general description of the steady component of ocean circulation (defined to be the temporal mean) has changed surprisingly little since World War II. In contrast, the view of temporal variability has undergone a major paradigm shift over the subsequent half century. Although eddy-like characteristics of ocean currents were known to exist even by

Maury (1855), it was difficult to distinguish unresolved variability from measurement errors. Multiship surveys and repeated hydrographic surveys conducted beginning in the 1950s and moored current meter and surface drifter measurements beginning in the 1960s revealed considerable spatial structure and temporal variability that did not support the view of ocean currents as simple and large scale. Much of modern oceanographic research has focused on understanding the nature of the rich spatial and temporal variability through a proliferation of new measuring and modeling techniques. There has been a growing recognition of the importance of short space- and time-scale variability (turbulence) to the large-scale circulation, momentum transport, and heat transport and to the distribution of chemical and biological properties.

Along with the rapid technological and theoretical developments over the past half century, oceanography has become progressively more specialized. It is no longer possible to summarize adequately the status of all disciplines of oceanography in a single book. Indeed, it is very difficult to summarize even a single discipline in one book. An excellent perspective on the post-World War II evolution of physical oceanography has been published by Warren and Wunsch (1981). A more popularized summary of several aspects of physical oceanography can be found in the Summer 1992 issue of *Oceanus* (Vol. 35(2)), which is dedicated to physical oceanography; dedicated issues on the other disciplines of oceanography can be found in the other 1992 issues of the magazine. A précis of physical oceanography is given in Chapter 1 of a National Research Council (NRC) report (NRC, 1988); also see NRC (1992b) for a state-of-the-science overview of all of oceanography.

In simple terms, physical oceanography can be defined as the study of the physics of the circulation of the ocean on all space and time scales. Research in physical oceanography includes studies of the details of turbulent mixing on scales of millimeters, the propagation of surface and internal waves with scales of centimeters to hundreds of meters, the dynamics of wind-forced and thermohaline-driven ocean currents (see, e.g., NRC, 1992b) on scales of kilometers to thousands of kilometers, and the transfers of momentum, heat, and salt within the ocean and across the air-sea interface. Because of the pressing importance of questions about global warming, there has been an increasing emphasis in recent years on the role of the ocean in the global climate. This has led to a quest for general understanding of the dynamics and long-term evolution of the coupled ocean-atmosphere system (see, e.g., Gill, 1982) and its interactions with the land, cryosphere, and biosphere. The need to quantify and forecast natural and anthropogenic changes in weather patterns and global climate, on the one hand, and the emergence of more easily accessible supercomputing power, satellite remote sensing, and other instrumentation technologies, on the other hand, are factors determining the direction of present and near-future research in physical oceanography.

Computer models of large-scale ocean circulation and ocean-atmosphere coupling, of biogeochemical cycles, and of the global budgets of carbon dioxide and other greenhouse gases are becoming the desired results of much of present research. The input data for such models have intrinsic shortcomings because of concerns about data quality and coverage (in space and time). Much effort must therefore be devoted to improving the interpretation of measured quantities and their subsequent use in computer models. The constraints may be due to limited spatial and temporal resolution of the measurements of the observed fields, limited accuracy of the measured quantities, gaps in the data records, short data records, or

propagation of errors through different levels of data processing and analysis. As a result, the technological innovations available do not guarantee success unless considerable progress is made in utilizing the available data. This will necessarily involve the use of sophisticated statistical techniques for a wide variety of purposes, as summarized in this report. Collaborative research involving statisticians and physical oceanographers is desirable to fuel such progress and improvements.

To provide statisticians with a brief sketch of the physical oceanographic community, the panel includes a few demographic items. It is not aware of any detailed demographic studies. The membership of the Ocean Sciences Section of the American Geophysical Union probably provides a fair representation of the community. In 1991, the section's total membership was 4791, 84 percent of whom were regular members and 16 percent of whom were student members. About one-fourth of this membership was foreign. Of the remaining members, it is not known what percentage are actively involved in research, but the number is probably less than half. The total membership is certainly dominated by physical oceanographers; it also includes a substantial number of chemical oceanographers and smaller numbers of biological and geological oceanographers, most of whom are members of other professional societies. About a dozen U.S. universities offer graduate programs in physical oceanography. There are two civilian federal government oceanographic laboratories and several U.S. Navy-supported research and development laboratories involved in open-ocean physical oceanographic research. Private industry employs a relatively small fraction of the physical oceanographic community.

Most physical oceanographic research is published in the six primary journals in the field: *Journal of Physical Oceanography*, *Journal of Geophysical Research-Oceans*, *Journal of Marine Research*, *Deep-Sea Research*, *Progress in Oceanography*, and *Journal of Atmospheric and Oceanic Technology*. Fundamental results frequently appear in the *Journal of Fluid Mechanics*. Significant advances in physical oceanographic research are occasionally published in *Science*, *Nature*, and *Geophysical Research Letters*. Overviews of physical oceanographic research written for less specialized audiences are often published in *Oceanography Magazine* and *Oceanus*.

OCEANOGRAPHIC MODELING, DATA, AND NOISE

The Many Meanings of the Term "Model"

The term "model" has a variety of usages in oceanography, depending on the context. It can refer to modeling of data by statistical methods (e.g., curve fitting of one-dimensional data, surface fitting of multi-dimensional data, correlation and regression analysis, modeling of probability distributions, and so on). More typically, however, the term "model" connotes physical modeling on the basis of mathematical equations that govern fluid motion, mass conservation, heat conservation, and conservation of salt or other chemical tracers. Physical models range from purely analytical (i.e., explicitly solvable in closed form) to numerical (i.e., solvable on a computer), depending on the degree of approximation of the complete mathematical equations adopted. An introduction to the equations of fluid motion in the

rotating reference frame of Earth can be found in Pond and Pickard (1983); a more advanced discussion can be found in Pedlosky (1987) or Stern (1975). A brief overview is given here.

The vector equation for momentum conservation based on Newton's Second Law that relates the acceleration of a fluid parcel to the forces acting on the parcel is

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + 2\boldsymbol{\Omega} \times \mathbf{v} = g - \frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{v}, \tag{1.1}$$

where $\mathbf{v}$ is the three-dimensional vector velocity, ∇ is the vector gradient operator along the x, y, and z coordinate axes with respective velocity components u, v, and w, $\boldsymbol{\Omega}$ is the angular velocity vector of the rotation of Earth, g is the gravitational acceleration, ρ is the water density, p is pressure, and ν is the molecular viscosity. The three components of this vector equation are referred to as the Navier-Stokes (N-S) equations, in honor of the physicist Claude L. M. H. Navier (1785-1836) and the mathematician Sir George Gabriel Stokes (1819-1903), who first formulated the molecular friction force in terms of the second derivatives of velocity along each of the three coordinate axes.

The unknown quantities in the N-S equations are density, pressure, and the three components of velocity. Two additional equations are thus necessary to solve for the five unknowns. The first of these is the mass conservation equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \tag{1.2a}$$

also known as the continuity equation. Seawater can generally be considered to be incompressible (i.e., the so-called total derivative $\partial\rho/\partial t + \mathbf{v} \cdot \nabla \rho$, corresponding to the rate of change of density following a fluid parcel, is zero), in which case the continuity equation reduces to

$$\nabla \cdot \mathbf{v} = 0. \tag{1.2b}$$

The other equation necessary to solve for the five unknowns is the equation of state relating density to temperature T, salinity S, and pressure,

$$\rho = \rho(T,S,p). \tag{1.3}$$

This empirical relationship is based on laboratory studies of seawater. The dependence of ρ on T and S requires the addition of two more equations governing the conservation of T and S. These equations have the form

$$\frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C = \kappa_C \nabla^2 C + Q_C, \tag{1.4}$$

where C could be either temperature or salt concentration, κ_C is the molecular diffusivity for C (analogous to the molecular viscosity ν in the N-S equations), and Q_C is a source or sink term to account for effects of heating and cooling. A source term is not necessary for salinity since all processes affecting salinity occur at boundaries (surface evaporation and precipitation, river runoff, freezing, and melting), and therefore enter the problem as boundary conditions. Temperature is also usually treated as a boundary condition, although, in a strict sense, the effects of solar heating can penetrate below the ocean surface.

In total, then, there are seven equations for the seven unknowns u, v, w, p, ρ, T, and S. These equations must be solved subject to boundary conditions of no normal flow at material surfaces (the ocean bottom and lateral boundaries), as well as boundary conditions for the normal and tangential components of forces at the boundaries (e.g., surface wind stress, bottom drag, lateral drag, and atmospheric pressure forcing) and buoyancy fluxes (heat and salt) across the air-sea interface and at coastal boundaries. The equations themselves are deterministic in the sense that a particular solution is obtained for a given specification of the boundary and initial conditions. However, the boundary and initial conditions have a random character, which imparts a randomness in the physical modeling.

It is noteworthy that many of the methods used to determine the ocean circulation are based on measurements of various natural and anthropogenic chemical tracers. Examples include oxygen, carbon dioxide, silicate, and tritium. The concentrations of these tracers are coupled to the dynamic variables of the equations of motion (1.1) and (1.2a) (or (1.2b)) through conservation equations with exactly the form (1.4), with the term Q_C corresponding to sources or sinks of the chemical tracer of interest. These tracers are used to infer indirectly the direction and, to some extent, the speed of deep ocean circulation where mean velocities are often too small to be measured directly.

The equations of motion apply to the instantaneous velocity of the fluid. However, the nonlinear terms in the momentum equation (1.1) give rise to turbulent variability that is characteristically irregular in space and time. Because of this nonlinearity and the large range of spatial scales over which the ocean is energetic, it is not practical to solve the above equations explicitly. In particular, it is not possible to measure, and hence specify, the boundary and initial conditions at very fine spatial and temporal resolution. This, in effect, introduces additional noise-like or random character to the physical equations. The usual approach to addressing the turbulent character of oceanic variability is to parametrize the effects of turbulence in terms of large-scale observable quantities (typically the mean flow and its derivatives). As a consequence of the neglect of the detailed dynamics on small scales, the parametrized physical equations pertain to averages of the random dynamic variables. The simplest and most commonly used approach is to replace molecular viscosity ν and diffusivity κ_C with "eddy" or "turbulent" viscosity and diffusivity (also referred to as effective diffusion or mixing coefficients), as first suggested by Taylor (1915). The turbulent coefficients serve the same function as molecular coefficients but are much larger in magnitude to account for the effects of eddies smaller than those explicitly represented

within the model. These eddies transport momentum and chemical properties much more rapidly than does molecular diffusion. Horizontal mixing is about 10 orders of magnitude larger than molecular diffusion. Because vertical density stratification in the ocean inhibits vertical mixing, vertical mixing is only about 2 orders of magnitude larger than diffusion.

The detailed specification of turbulent mixing is not well understood because, unlike molecular diffusion, which is an intrinsic property of the fluid, turbulent mixing varies spatially and temporally and depends on the flow itself. Moreover, the particular choice of turbulent mixing coefficient depends critically on the spatial scales represented within the model. From coarsely spaced observations, it is even possible for turbulent transport to be counter-gradient (i.e., effectively a negative turbulent mixing coefficient, corresponding to energy transfer from eddies to the mean flow; see Starr, 1968). Such a situation is clearly nonphysical, and the turbulent mixing coefficient would presumably be non-negative with sufficiently close sample spacing.

The equations of motion (1.1)-(1.4) (referred to as primitive equations) are very complex and are therefore not solvable in exact form. Various simplifications of the complete equations are employed in order to gain insight into the dynamics of fluid motion. A brief overview is given here; a more detailed summary can be found in Holland (1977). One class of simplifications concerns the treatment of vertical density stratification. The simplest models, referred to as barotropic models, consider the fluid density to be homogeneous. Next in complexity are layered models that divide the ocean into two or more distinct layers, in each of which the fluid density is considered homogeneous. The most complex models consider the fluid to be continuously stratified. Although a barotropic approximation is clearly unrealistic, many circulation aspects can be successfully modeled without the need for the more complex baroclinic layered or continuously stratified models.

For both barotropic and baroclinic models, various approximations are employed to simplify the equations of motion. The simplest model is the geostrophic approximation, which neglects the nonlinear and acceleration (i.e., time-dependent) terms. The resulting steady-state, linearized equations can be solved analytically, and the geostrophic solution is surprisingly successful at describing the large-scale aspects of the circulation. The next level of complexity includes the acceleration term, which permits analytical wave solutions. Depending on the scales of interest, these waves can range from short capillary waves (wavelengths of millimeters) for which the restoring force is surface tension, to surface and internal gravity waves (wavelengths of tens of centimeters to hundreds of meters) for which the restoring force is gravity, to very long wavelength (tens to hundreds of kilometers) Kelvin or quasi-geostrophic Rossby waves, which arise from the restoring force provided by the latitudinal variation of the local vertical component of Earth's angular velocity vector or horizontal gradients of bottom topography. The large-scale waves are the dynamical mechanism by which the large-scale circulation adjusts to time-dependent forcing such as the stress exerted by the wind blowing over the surface of the ocean.

Although very illuminating, linear models of ocean circulation are not capable of producing accurate representations of detailed aspects of the circulation. In particular, the short spatial scales of many of the interesting features of the circulation (e.g., jetlike currents such as the Gulf Stream) result in strong gradients in the velocity field, which elevates the magnitude of the nonlinear terms to a level comparable to that of other terms in the

equations of motion. More complex classes of physical models thus include nonlinear effects. Analytical solutions are still possible for weakly nonlinear approximations and for a few special cases of strongly nonlinear approximations of the equations of motion. Numerical methods using a computer are necessary for more general solutions.

Numerical models can be classified as either process-oriented (also referred to as mechanistic) or simulation models. Process-oriented models simplify the ocean basin geometry in order to focus on the physics of specific term balances in the mathematical equations. Simulation models attempt to represent the basin geometry more accurately and to reproduce or predict some aspects of the actual circulation for comparison with observations. Numerical solutions to the equations of motion are obtained on a space-time grid by approximating the derivatives in the equations by finite differences or by the use of Fourier transform techniques. At each grid point, solutions are obtained by stepping forward in time from the initial conditions according to the mathematical equations governing the fluid motion (e.g., O'Brien, 1986; Haltiner and Williams, 1980; NRC, 1984).

Computational models of the climate, especially coupled ocean-atmosphere models, are being used to produce estimates of the climate changes to be expected to result from changes in radiative forcing. Although deterministic, these models are sufficiently chaotic to show variability that is in many respects similar to that observed in the climate of the real world. Thus, the analysis of model output and comparison with data (see Chapter 7), especially to detect trends, raises serious statistical questions.

The accuracy of a numerical solution depends critically on the spatial resolution of the grid and on the size of the time step, as well as on the particular parametrizations of the turbulent viscosity and specifications of the boundary and initial conditions. There are thus many ways in which the mathematical equations governing the physics of the ocean can be solved numerically. In general, the most accurate simulations require very fine grid spacing and short time steps. In practice, spatial and temporal resolutions are limited by available computer time and memory allocation. Disk storage capacity can also present a problem since the volume of model output can be very large. As discussed in Chapter 5, physical oceanographic research would benefit greatly from improved methods of visualization to examine the four-dimensional output of numerical models of ocean circulation.

Besides the difficulties associated with the subjective natures of the choice of grid resolution, parametrization of turbulent viscosity, and the problem of availability of computer resources, another major issue in physical modeling of the ocean is assessment of the accuracy of the solution. Due to the underlying chaotic nature of ocean circulation (e.g., Ridderinkhof and Zimmerman, 1992), to numerical inaccuracies, and to inaccuracies in the specifications of boundary and initial conditions, numerical simulations can be expected to diverge fairly quickly from the actual circulation. One of the challenges of modern physical oceanography is development of techniques for comparing simulations from different numerical models with each other and with one or more independent observational data sets in order to evaluate the relative accuracies of various model simulations. It is unlikely that numerical simulations can ever be expected to exactly depict the actual circulation. There is currently no general agreement about what aspect of model simulation is most important. For example, one measure of the accuracy of a model is how well it represents the mean circulation. Another measure of accuracy is how well higher-order statistics of the flow field

are reproduced (e.g., the variance of a particular variable or the covariance between two variables). As discussed in Chapter 7, data and model cross-comparison is another area in which the field of statistics may be able to make important contributions.

It is noteworthy that, in contrast to physical modeling of atmospheric circulation, the detailed evolution of the actual ocean circulation is very poorly known because of a lack of observations. Global coverage of the ocean can be obtained only from satellite observations, but these are nonsynoptic (i.e., not simultaneous at all locations over Earth) and sample only surface conditions. Sparsely distributed in situ measurements or physical modeling (or both) are necessary to extrapolate the surface measurements from satellites to infer the ocean circulation at depth. Much of the present emphasis in physical modeling of the ocean is directed at developing methods of assimilating available observations (especially satellite observations) into the model solution at regularly or irregularly spaced time steps using statistical estimation, Kalman filtering, and generalized inverse techniques. Such methods have been in use in meteorology for some time. Recent reviews of oceanographic applications of data assimilation can be found in Ghil and Malanotte-Rizzoli (1991) and Bennett (1992). Successful assimilation of available data preserves some degree of similarity between numerical solutions and the actual circulation.

Diverse Definitions of the Term "Data"

Clarification is in order regarding oceanographic usage of the term "data." In the field of physical oceanography, the term is used more liberally than in some other fields of science. The intent here is not to justify oceanographic use (or misuse) of the term, but rather to clarify the standard oceanographic jargon and the usage elsewhere in this report. Unlike measurements in some fields of science, few, if any, oceanographic measurements are direct. The quantity of interest is typically sensed electronically as a voltage drop, the number of frequency oscillations of a quartz crystal, the number of rotations of a rotor, or a count of some other sort. These counts must be converted to the geophysical quantity that is of interest by a hierarchy of transformations, some of which may be nonlinear or irreversible. These transformations are often empirically based and could benefit from improved statistical formulations.

At each level of transformation, the output of the previous transformation becomes the input for analysis or for a higher level of transformation. This input is then generally referred to as "data" and is typically treated as if all previous levels of transformation have been done correctly. In this context, then, even the output of a numerical ocean model forced by wind fields derived from in situ or satellite observations can be, and sometimes is, referred to as "data" by an investigator interested in analyzing the model output to study ocean dynamics. An important element of these multiple levels of transformation is that it becomes progressively more difficult, and sometimes even impossible, to quantify uncertainties in the output product.

Multiple levels of transformation are characteristic of all oceanographic data but are especially pronounced for satellite data. In an effort to distinguish between different types of "data," NASA defined a hierarchy of data levels in the early 1980s (see, e.g., Arvidson et

al., 1986; Dutton, 1989). The same definitions have subsequently been used for in situ observations, although some definitions of data level are not appropriate for some types of in situ data. A summary of the data levels follows:

Level 0: Raw instrument data at original resolution, time ordered, with any duplicates removed. For satellite observations, this level of data consists of the bits (possibly compressed for transmission) telemetered from the satellite to a ground receiving station, corrected for any telemetry errors. For in situ observations, this level of data might consist of volts or counts of some other type. Level-0 data are sometimes referred to as experimental data.

Level 1A: Reformatted or reversibly transformed level-0 data, located to a coordinate system (e.g., time, latitude, longitude, depth) and packaged with needed ancillary, engineering, and auxiliary data. Instrument counts from level-0 data have been converted to engineering units in level-1A data. In the case of in situ data, level-0 and level-1A may be the same.

Level 1B: Irreversibly transformed values of the instrument measurements. For satellite observations, this might consist of calibrated microwave antenna temperatures, infrared or visible radiances, or microwave normalized radar cross sections. For in situ observations, this level of data is typically the geophysical parameter of interest. In some cases, the data might be resampled to a new grid.

Level 2: Geophysical parameters at the measurement time and location. For satellite observations, level-2 data are obtained from a model function (typically derived empirically from some statistical analysis) applied to the level-1B data. For in situ observations, level-2 data may be the level-1B geophysical parameters corrected for any systematic errors or calibration adjustments (typically determined empirically from some statistical analysis).

Level 3: Geophysical parameters resampled onto a regularly spaced spatial, temporal, or space-time grid by some sort of averaging or interpolation.

Level 4 and above: No set definitions, but generally refer to higher-level processing. An example would be a map of some statistical quantity such as the mean value or standard deviation of a lower-level data quantity. Another example would be higher-level wind fields derived from gridded fields of surface wind velocity (e.g., wind stress or the curl of the wind stress, both of which are used for studies of wind-forced ocean circulation). An extreme example is the output of a numerical ocean circulation model forced by wind fields derived from a level-3 wind product.

Specific examples serve to clarify the need for multiple data-level definitions in oceanography. Virtually any oceanographic measurement could serve as an adequate example for this purpose. The following two examples (one a satellite measurement and the other an in situ measurement) were chosen rather arbitrarily:

• **Example 1: Near-surface vector winds estimated by a satellite radar scatterometer.** The basic quantity measured by the scatterometer is the power of the radar return. The measured return power is digitized, compressed, and telemetered to a ground receiving station along with a variety of necessary ancillary information (e.g., orbit altitude, satellite attitude, temperatures of the electronic components, and so on). The telemetry "data" are uncompressed and converted to engineering units "data" in ground-based processing. A quantity referred to as the normalized radar cross section (NRCS) is derived from the measured return power by normalizing by the power of the transmitted signal along with any necessary calibration adjustments determined from prelaunch calibration or from the ancillary information. Estimates of vector winds are constructed from NRCS "data" from two or more antenna look angles, collocated at approximately the same location on the sea surface. This requires both an empirically derived model function and a statistical method for solving the overdetermined problem of inverting the model function in a manner that is consistent with the noisy NRCS "data." The result at this stage is individual vector wind "data" at the measurement locations. Most oceanographic applications of scatterometer observations require gridded fields of vector winds or some higher-level wind product derived from Earth-located individual vector wind "data." These fields are obtained by space-time averaging or interpolation and are generally referred to as "data" by investigators who analyze the wind fields or use them to force ocean circulation models.

• **Example 2: Measurements of temperature and salinity by a conductivity-temperature-depth (CTD) profiler.** A CTD (e.g., see p. 389 in Dickey, 1991) is lowered through the water column on a cable. Variations in voltage associated with changes in temperature and conductivity are measured at a high frequency from two separate sensors (a thermistor and a conductivity probe). These engineering unit "data" are converted to temperature and conductivity "data" through simple algorithms. The conductivity of seawater is a function of both temperature and salinity. Temperature effects are much greater than salinity effects and must therefore be removed from the conductivity measurements in order to estimate salinity. However, the response time of the thermistor measurements of temperature alone is much longer than the response time of the conductivity probe because of thermal inertia of the thermistor. This difference in response time must be accounted for when using the thermistor measurements to remove the temperature component of conductivity variations. Salinity "data" compatible with the thermistor measurements are usually obtained by applying a low-pass filtering algorithm to effectively slow down the response of the conductivity probe. The resulting temperature and salinity "data" at closely spaced vertical intervals usually are then bin averaged and processed to reduce the data volume. It is also necessary to adjust the salinity and, to a lesser extent, the temperature estimates to account for periodic recalibrations of the two sensors. The resulting vertical profiles of temperature and salinity "data" are useful for many oceanographic applications. Some applications require further processing of the temperature and salinity "data" to derive density, thereby yielding a vertical

profile of density. The density "data" may then be integrated vertically to estimate the so-called steric height of the sea surface (or any other isobaric surface) relative to an arbitrary reference level. Density profiles, steric height, and other higher-level "data" derived from the CTD temperature and salinity "data" are typically used to construct vertical sections or horizontal maps of the quantity of interest. These sections or maps are often referred to as "data" by investigators who analyze them or use them to force ocean circulation models or to verify ocean model output.

Because of the multiple scales characteristic of both spatial and temporal variability in the ocean as discussed in Chapter 2, oceanographic data are commonly undersampled in several respects. One problem is aliasing that arises as a consequence of practical considerations that often limit the sampling to spatial or temporal intervals that are longer than the shortest energetic space and time scales of variability of the quantity being measured. For example, time series constructed from satellite observations are limited by the time interval between repeated satellite orbits over a given location. As another example, temperature measurements from an instrument lowered through the water column are sampled discretely at a fixed rate that often does not adequately resolve variations on the vertical scales of millimeters to centimeters that are important to turbulent mixing. As a third example, lines of vertical profiles of temperature and salinity along hydrographic sections across an ocean basin are sometimes not sampled sufficiently often along the ship track to resolve the energetic 10- to 50-km mesoscale variability that is superimposed on the larger-scale 100- to 1000-km variability that may be the primary signal of interest. The degree to which aliasing affects oceanographic data depends on the energy of the unresolved variability, be it of high frequency or short spatial scale, compared with the energy of the oceanographic signal of interest for the particular application of the data.

Another common problem is the limited spatial or temporal resolution inherent in many oceanographic measurements because of limitations of the measurement process. For example, satellite data generally consist of instantaneous measurements effectively averaged over a relatively large spatial "footprint." As another example, current meter measurements often consist of a time series of successive time averages at a fixed location. In some cases, the spatial or temporal averaging obscures signals in the quantity being measured that might be of interest for some studies. In others, time series may be uncomfortably short, important concomitant variables may not have been measured, and other factors may be contaminating the records. For example, a change in instrumentation or recording sites can limit the amount of useful information contained in a data set. There may be gaps in the records and the raw (level-0) data may not be readily accessible.

Such processes often generate measurements that violate the assumptions of the simplest statistical theory; i.e., the data are typically not independent, are not identically distributed, are not stationary, are non-Gaussian, or some combination. Especially problematic in this regard is serial dependence, which occurs at least to some extent in nearly all temporal oceanographic data.

Collected data can involve a sampling problem because of the fundamentally "red" spectral characteristics of ocean variability (i.e., the predominance of energy at the lowest frequencies). Most oceanographic data records are not long enough to resolve all of the

time scales of variability of the quantity of interest. This limits the frequency and wavenumber resolution of the measurements and the number of independent realizations of the physical process of interest. For example, the El Niño phenomenon that affects much of the ocean and the overlying atmosphere has a time scale of 3 to 5 years (cf., Ropelewski, 1992). Even a 30-year record (which is unusually long for physical oceanographic data) only resolves 6 to 10 realizations of this process, resulting in limited degrees of freedom for inferences about cause and effect (see, e.g., Davis, 1977; Chelton, 1983; Thiebaux and Zwiers, 1984; Barnett and Hasselman, 1979).

An important example of unresolved variability is the secular trend of sea level rise (see, for instance, NRC, 1990b) associated with global warming (see also, Baggeroer and Munk, 1992). The study of oceanic sea levels is further complicated by there being very few long data records, and by the existence of other poorly understood signals in the data (for example, glacial rebound effects). The data also include long-period signals, such as the 18.6-year lunar tide. The processes responsible for changes in sea level need to be understood, and especially in their relation to possible global warming. If the oceans were to warm, thermal expansion of seawater would be reflected in increased sea levels, with obvious effects on human activity.

Coupled with the problem of limited record length is the problem that many oceanographic signals of interest are intermittent (i.e., non-stationary or non-homogeneous). For example, turbulent mixing in the ocean generally occurs in sudden bursts and spatially irregular patches. Another example is the energetic wind events such as storms that vigorously force the ocean but occur only intermittently at a given location. As a consequence, it is difficult to characterize the statistics of ocean variability. For some purposes, it is the intermittent events that are of interest. In other applications, energetic intermittent events might be considered nuisances that can skew the sample statistics (e.g., the mean value or variance) that may be of interest. Techniques for analysis of non-Gaussian data (see Chapter 8) or estimation of robust statistics are therefore needed for many analyses of oceanographic data.

These data provide the statistician and data analyst with many challenges. For example, work needs to be done on multivariate transfer functions, particularly with mixed spectra. Data such as these often contain both large deterministic effects and periodic terms plus a non-deterministic part. This can cause serious problems of estimation. Short multivariate series for which the number of series is greater than the number of temporal observations provide a particular challenge because any standard estimate of the spectral matrix is singular. An example of this type of problem is spatial temperature series for which the assumption of spatial homogeneity is obviously not appropriate, but, at least in some regions, spatial continuity might be reasonable. In many of these instances, estimates of uncertainty are inadequate or are completely lacking.

Low Noise Is Good Noise

Oceanographic measurements often suffer from low signal-to-noise ratio, in some cases because the signal of interest has much smaller energy than other geophysical signals

in the data. For example, the sea level rise from global warming is much smaller than the energetic sea level variations of other oceanographic and non-oceanographic origin (see Chelton and Enfield, 1986). As another example, the visible radiances measured from a satellite for estimation of ocean chlorophyll concentration and investigation of the role of the ocean in the global carbon budget are dominated by atmospheric contamination from the scattering of sunlight from aerosol particles and atmospheric molecules; only about 20 percent of the measured radiances originate from the ocean (Gordon and Castano, 1987). A low signal-to-noise ratio may also arise because of the short record lengths typical of oceanographic data compared with the time scales of the signal of interest. Quantifying the signal-to-noise ratio and the auto- and cross-covariance functions of the signal and noise are important challenges in physical oceanography. A particularly difficult problem arises because of the fact that low-frequency calibration drifts in the measuring devices are often as large in magnitude as the low-frequency signal of interest. For example, estimation of sea level rise from global warming is complicated by vertical crustal motion in the vicinity of many ocean tide gauges. As another example, estimation of low-frequency variations in bottom pressure is complicated by electronic drifts in the pressure gauge measurements.

Because of the variety of sampling problems inherent in oceanographic data, the term "noise" is often used to refer to more than just the measurement error associated with inaccuracies in the observations. Inadequately resolved contributions to a measurement from geophysical variability of the quantity of interest are generally referred to as "geophysical noise." As discussed above, such unresolved geophysical variability can arise from use of a discrete sample interval (aliasing), from inherent spatial or temporal smoothing in the measurement (limited resolution), from finite record length (limited frequency or wavenumber resolution), from intermittency of energetic signals other than those of primary interest, or from low signal energy compared with the geophysical noise of other processes affecting the measured quantity. Although such geophysical noise is fundamentally different from that due to measurement errors, it has exactly the same effect as measurement errors from the point of view of data analyses. When there is a low signal-to-noise ratio, extraction of the signal of interest is especially difficult because typically the measurement noise and geophysical noise in the data are serially correlated.

2

STATISTICAL ISSUES IN THE MULTIPLE-SCALE VARIABILITY OF OCEANOGRAPHIC FIELDS

OCEANOGRAPHIC VARIABILITY

Oceanographic fields and processes possess certain features that are not commonly encountered in some other areas of science and engineering. One of these is a wide range of scales (wavenumbers and frequencies) in which observed fields exhibit spatial and temporal variation. In other words, a "typical" time (space) scale is absent, and there exists a broad band of frequencies (wavenumbers) of roughly equal importance. This is the reason for the term "multiple-scale variability." Oceanographic processes include coupling across a large range of scales (i.e., nonlocal interactions) and linkage between a number of factors of different nature. In Figure 2.1 (from Dickey, 1990, 1991), typical spatial and temporal scales of some oceanographic processes are sketched.

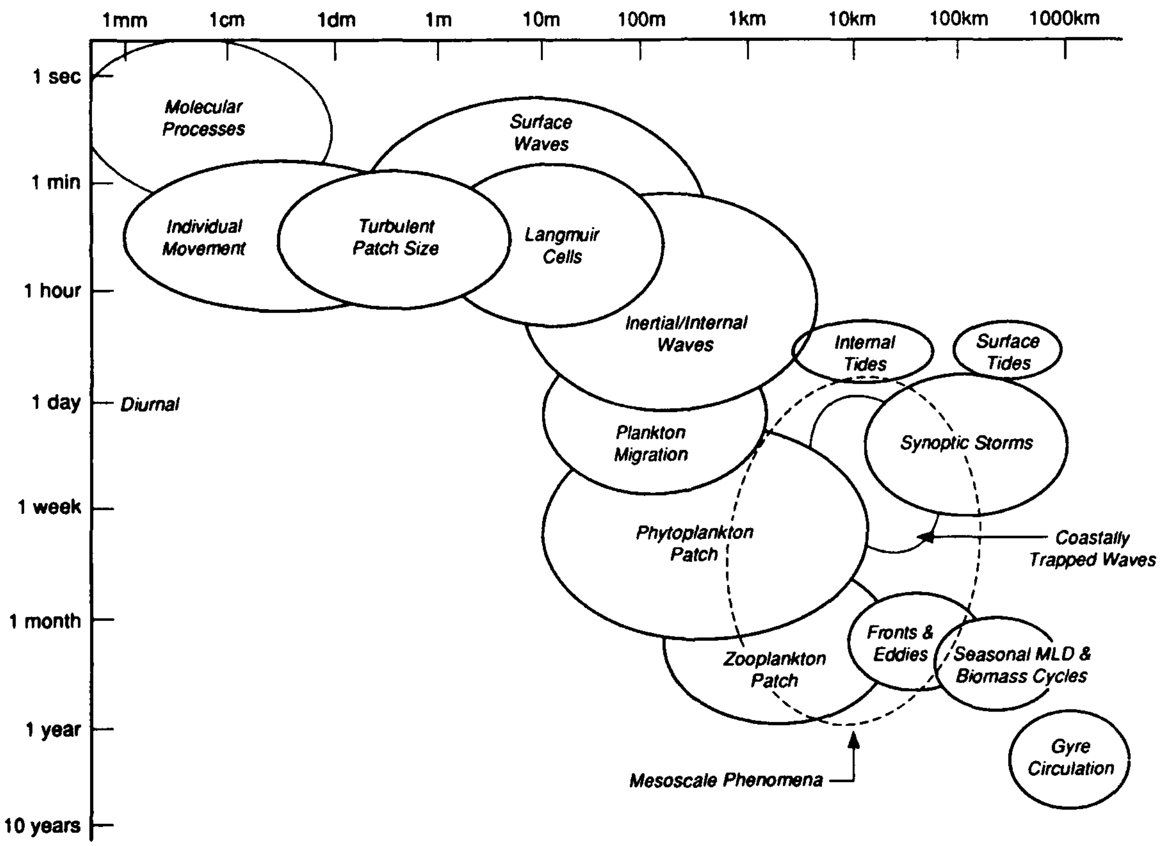

FIGURE 2.1 A schematic diagram illustrating the relevant time and space scales of several physical and biological processes important to the physics and ecosystem of the upper ocean. Reprinted from Dickey (1990, 1991) with permission.

From the statistical standpoint, a random field is a stochastic process with multidimensional parameters (e.g., time and position) or a more complicated parameter such as a function. The fields of primary interest have four parameters: one dimension of time and three dimensions of space. Examples of such time-varying fields include fluid velocity, pressure, water density, temperature, and salinity. Fields with only two spatial dimensions include sea surface height (sea level), wind velocity and wind stress at the surface, sea surface temperature (SST), ocean color, and sea ice. Wavenumber spectra of these fields are usually very broad, covering several decades of wavenumbers (e.g., Fu, 1983; Freilich and Chelton, 1986), and the spectral density function can be approximated by a power law. Characteristic values of exponents in the power laws indicate a fractal regime in the geometry of the fields. For instance, the sea surface elevation field, for scales related to wind-generated surface gravity waves (from a decimeter to several hundred meters), is characterized by a two-dimensional wavenumber spectrum that falls off roughly as $k^{-7/2}$. This corresponds to a cascade pattern in surface topography (a hierarchy of randomly superimposed waves with decreasing amplitude and wavelength). A characteristic property of this field is its statistical self-affinity (Glazman and Weichman, 1989). The corresponding Hausdorff dimension, for an assumed Gaussian distribution, is 2.25.

The fluid velocity field, whose kinetic energy spectrum is characterized by $k^{-5/3}$, exhibits a Hausdorff dimension of 2.666. A typical geometrical feature of such fields is a hierarchy of eddies. Such cascade patterns in a field's geometry are related to the cascade nature of the energy transfer along the spectrum through nonlinear interactions among different scales of fluid motion. Other physical quantities, e.g., momentum, enstrophy (i.e., half the square of vorticity), and wave action, may also be transferred either up or down the spectrum. The spectral cascades of these quantities are not *necessarily conservative:* interactions between different oceanographic fields (occurring within certain limited ranges of scales — the "generation and dissipation subranges" — and resulting in energy and momentum exchange) provide energy sources or sinks in various spectral bands. For instance, at meter scales wind provides the energy input into surface gravity waves that in turn exchange momentum and energy with larger-scale motions (e.g., mesoscale eddies, Langmuir circulations, internal waves). Mesoscale oceanic eddies are caused by the barotropic instability of basin-scale currents. Seasonal heating and cooling of the ocean surface causes convection and vertical mixing, while differential (across the oceanic basins) heating, evaporation, precipitation, and ice melting cause density-driven currents. Ocean circulation on basin scales is caused by large-scale curl of the wind stress. This multiplicity of the energy sources and sinks and the interactions between different scales and individual components of ocean dynamics are responsible for the extreme complexity of patterns of ocean circulation, sea surface temperature, sea level, and so on as observed both in satellite images and in highly complicated trajectories of free-drifting floats. Apparently, the interaction of motions with different scales implies statistical dependence between corresponding Fourier components or between corresponding eigenvectors in the empirical orthogonal functions (EOF) series (Karhunen-Loeve expansion; see, e.g., Lorenz, 1956; Davis, 1976; Preisendorfer, 1988). Identifying and accounting for such correlations in statistical models are important problems of oceanographic data analysis.

The difficulties mentioned above need not defeat efforts to understand ocean dynamics. In contrast to economics, demography, biology, and many other fields, physical oceanography is based on the comparatively reliable and universal quantitative physical models summarized in Chapter 1.

Initial and boundary conditions complete the formulation of specific oceanographic problems. Since the boundary conditions (e.g., the distribution of wind stress over the sea surface) and the coefficients in the equations (e.g., ocean current velocities in the heat-transfer equations) are intrinsically random, oceanographic problems are actually those for stochastic partial differential equations (SPDEs). Many of the issues related to SPDEs are also encountered in analysis of oceanographic observations. These include, for instance, the impact of subscale (microscopic) motions on the (macroscopic) behavior of the mean fields (analogous to the dependence of measured quantities on the spatial, temporal, or spatio-temporal resolution of a measuring technique). On a more fundamental level, the justification of the "macroscopic" equations remains a difficult problem.

These problems that present opportunities for statisticians are also central to eventually understanding the structure of turbulent flow. Turbulent fields of fluid velocity, pressure, and temperature are highly inhomogeneous and include compact regions where these fields or their spatial derivatives attain extreme values. Regions with large fluid velocity gradients are particularly important, because most dissipation of the mechanical energy into heat occurs in these localized regions. Due to an irregular spatial and temporal distribution of such regions, the occurrences of extreme events are often referred to as intermittency. Intermittency becomes pronounced at high Reynolds numbers associated with the onset of turbulence. The Reynolds number is a measure of the relative importance of inertial forces in the fluid as compared to viscous forces (viz., it is the ratio of the inertia of fluid particles to the fluid's viscous friction). At high Reynolds numbers, when the inertia of fluid particles is no longer balanced by friction forces, particle trajectories become tremendously complicated. This results from the frictionless fluid particles having an unrestrained ability to continue their motion in whatever may be the direction they were launched (by some initial disturbance) or deflected (by interactions with neighboring particles). No matter how small the differences in initial directions and velocities between individual particles, their trajectories quickly diverge. An observer sees a highly chaotic pattern of flow, including intermittent events with particularly large velocity gradients. What is the probability structure of the dissipation field and related field gradients in a turbulent flow? No rigorous deductions based on the governing N-S equations have been reported, although a number of heuristic models have been proposed (e.g., Novikov and Stewart, 1964; Novikov, 1966; Yaglom, 1966; Mandelbrot, 1974).

SATELLITE OBSERVATIONS

Satellite instruments measure at different incidence angles the electromagnetic characteristics of the emitted radiation (passive instruments working in visible, infrared, and microwave ranges of the electromagnetic spectrum) and backscattered radar pulses (active instruments working in the microwave range) that come from the ocean surface. These

characteristics (e.g., the intensity of visible and infrared radiation at various wavelengths, radio brightness temperature, radar cross section, round-trip travel time of a reflected pulse, the shape of the pulse distorted by a random sea surface, and so on) are interpreted in terms of oceanographic parameters (pigment and chlorophyll-A concentrations, sea surface temperature, wind speed and direction at the surface, sea level height, and others). The interpretations are typically obtained from empirical algorithms based on incomplete or approximate physical models. For instance, empirical relationships based on a limited set of coincidental radar and buoy observations are routinely employed to derive wind speed from altimeter and scatterometer radar cross sections. Such relationships are called geophysical model functions (GMFs). The available GMFs are based on rather simple linear or nonlinear regression models, and considerable improvement might be possible in this area with the use of more advanced statistical methods.

Instrument footprint sizes, swath widths, and other characteristics of typical satellite instruments are summarized in Table 2.1. The footprint is a spot on the surface from which reflected or emitted radiation is collected by satellite antenna to produce the observed radar cross section, brightness temperature, and so on. Spatial coverage (which depends on swath width, footprint size, sampling rate, and satellite orbit geometry) varies from one instrument to another. The spatial sampling rate, i.e., the distance between individual satellite footprints, may cause aliasing of the data. Other factors leading to aliasing are the spatial separation of satellite orbits and the specific time interval between repeat tracks (see Figure 6.1 in Chapter 6). All these factors raise issues regarding correct interpretation of satellite measurements and their use in numerical models of ocean circulation. Spatial inhomogeneity of surface properties on scales within and beyond the footprint size, and these properties varying *nonlinearly along any direction within a footprint*, produce an appreciable dependence of satellite measurements upon the instrument employed. The case of wind speed measurements is most instructive. Wind speed maps for the same period of time but based on measurements by different satellite techniques exhibit appreciable differences—regardless of the fact that the root-mean-square measurement errors characterizing individual instruments are very similar. Pandey (1987) compared wind fields based on satellite scatterometer, altimeter, and microwave radiometer data and found that the discrepancy locally may exceed 2 m/s. Statistical distributions of wind velocities derived from different instruments can also differ.

Statistical models of oceanographic fields with prescribed statistical properties might prove useful for analysis of satellite and other measurements (e.g., Ropelewski, 1992). In Chapter 6, additional problems arising in connection with the spatial inhomogeneity, statistical anisotropy and intermittency observed in oceanographic fields are reviewed. Those include transferring (binning) the satellite-produced data onto geographic grids, filling gaps in the data, and interpolating, extrapolating, smoothing, and filtering the data.

ISSUES FOR STATISTICAL RESEARCH

There are important open questions associated with sampling at different rates: how does sampling at different rates relate to aliasing, and to interaction of processes occurring

TABLE 2.1 Characteristics of Satellite Microwave Instruments for Ocean Studies

Instrument and Its Main Features	Measured Electro-magnetic Parameters	Inferred Ocean Parameters	Footprint	Swath Width	Additional Information
Altimeter: sends short pulses at nadir incidence (13-GHz carrier frequency; TOPEX altimeter will also have a 5-GHz channel)	Travel time of a return pulse, radar cross section, shape of a return pulse	Sea level height Wind speed Significant wave height	Circular, 5- to 12-km diameter, depending on surface roughness	One-pixel diam. ≈ 10 km	Along-track pixel spacing: ~7 km. Distance between tracks at equator: ~150 km. 10 to 20 days exact repetition of all orbits
Scatterometer: sends short pulses in a range of incidence angles from 20 to 60 degrees, using both strictly horizontal (HH) and strictly vertical (VV) polar-izations; 14-GHz carrier frequency	A set of radar cross sections for each surface bin, at several azimuthal angles and polariza-tions	Wind speed Wind direction	Aspect ratio ~1:4. Major axis: 30 to 90 km depending on position within the swath, etc.	Two swaths 600 to 700 km each	Global coverage every 2 days
Synthetic aperture radar: high-spatial-resolution radar images of sea surface roughness distribu-tion for C, L, and X bands. Other bands have also been employed	Analog and/or digital matrices of radar cross section showing spatial varia-tion of surface roughness	Length and direction, or surface gravity and internal waves, wave number spectra of surface roughness spatial varia-tion, surface signatures of mesoscale eddies, fronts, current boundaries, sea ice, bathymetry	10- to 100-m linear size, depending on the mode frequency electromag-netic band, etc.	Hun-dreds of kilo-meters	Usually only regional coverage for selected locations
Special Sensor Microwave/Imager (SSM/I) with channels (GHz): 19.4 22.2 37.0 85.5	Radio brightness temperature	Characteristics of atmosphere (e.g., water content); surface wind speed, sea ice	Len-gth / Width (km) 70 45 60 40 38 30 16 14	1300 km	Almost total global coverage obtained every day
Scanning Multichannel Microwave Radiometer (SMMR) with channels (GHz): 37.0 21.0 18.0 10.7 6.6	Radio brightness temperature	Characteristics of atmosphere (e.g., water content); surface wind speed, sea surface temper-ature, sea ice	Len-gth / Width (km) 22 14 28 25 43 28 74 49 120 79	780 km	Almost total global coverage every 2 days

SOURCE: Courtesy of Roman Glazman, Jet Propulsion Laboratory, California Institute of Technology.

at different scales? What can and cannot be inferred about the continuous process within which sampling is done? These concerns also involve different types of estimates such as second- and higher-order spectral estimates, probability density estimates, and regression estimates. Such questions should be considered under the assumptions of both stationary and nonstationary processes. These problems are connected with those involving non-Gaussian observations (see Chapter 8). Suitably selected and designed multiscale wavelets may be helpful in this situation.

There are statistical research opportunities in modeling a random field given:

1. observational data representing averages over regions (pixels) of a given size (as determined, e.g., by a satellite footprint), and
2. observational data obtained by irregular sampling (spatial and temporal data gaps, etc.) of a random field.

An analysis of extrema of non-Gaussian fields is needed. It will depend partly on what one can say in the stationary case about the tails of the instantaneous distributions. Such an analysis will have both a probabilistic and a statistical aspect; i.e., given a nice probabilistic characterization, can some aspect of it be effectively estimated from data? Progress on these questions may also carry over to notions of intermittency. Specific issues for focus include:

1. analysis of asymptotics of extrema of a non-Gaussian field,
2. analysis of behavior of outlying observations in a case of non-Gaussian data, and
3. modeling of a random field with given statistics of extrema.

Additional issues and problems concerning non-Gaussian random fields and processes are listed at the end of Chapter 8.

3
LAGRANGIAN AND EULERIAN DATA AND MODELS

In the last two decades the use of Lagrangian (i.e., current-following) devices has become very popular in oceanography (for a review, see Davis, 1991a). Drifting buoys have been developed that can follow the ocean currents with good accuracy, moving either at the surface of the ocean or in the interior on surfaces of equal pressure or density. These drifting buoys are tracked acoustically or via satellite for extensive time after deployment (up to a year or more). They report their position at discrete times, with an interval that can vary from hours to days depending on the specific purpose of the measurements made. From these positions, an estimate of the horizontal velocity along the buoy trajectory can be made. In addition to their position, drifting buoys are often equipped to measure other physical quantities, such as temperature or pressure.

Data from drifting buoys are used both for understanding the dynamics of ocean circulation (e.g., Price and Rossby, 1982; Bower and Rossby, 1989) and for describing its statistical properties (e.g., Kraus and Boning, 1987; Figueroa and Olson, 1989). This chapter focuses on this second aspect. An appropriate statistical description of ocean circulation includes two main parts. One is the statistics of the velocity field, and the other is the statistical description of the transport mechanisms. The ocean plays a fundamental role in the transport of such quantities as heat, salinity, or chemical substances (both natural and anthropogenic) that are fundamental for environmental and climatic studies. Before going into the details of how the Lagrangian data are actually utilized to obtain the statistical information, it is useful to point out that there is a direct connection between Lagrangian trajectories and transport properties in a flow (e.g., Davis, 1983). This can be seen by considering the equation for the evolution of the concentration of a substance released and transported in an incompressible fluid: $(\nabla, \boldsymbol{u}) = 0$ (see, e.g., Pedlosky 1987). Assuming that the substance concentration is a scalar function $c(t, \boldsymbol{r})$, and that the substance does not interact with the flow while it is advected (i.e., it is a passive scalar, or "tracer"), the equation is

$$\partial_t c + (\boldsymbol{u}, \nabla) c = 0, \quad c(0,x) = c_0. \tag{3.1}$$

Note that equation (3.1) is the same as equation (1.4) of Chapter 1, except that the molecular diffusivity is neglected because here the concern is large-scale flows, and for simplicity no sources or sinks are considered. The solution of equation (3.1) by the method of characteristics takes the form

$$c(t,\boldsymbol{r}) = c_0(\boldsymbol{X}^{-1}(t,\boldsymbol{r})), \tag{3.2}$$

where $\boldsymbol{X}^{-1}$ is the inverse of the function $\boldsymbol{r} \longrightarrow \boldsymbol{X}(t, \boldsymbol{r})$ that represents the position reached at time t by a particle that was at $\boldsymbol{r}$ at $t = 0$.

From (3.2) one can calculate statistical moments of the concentration $c(t,\boldsymbol{r})$ by the formula

$$\langle c(t,r_1)c(t,r_2)\cdots c(t,r_p)\rangle =$$
$$\int_{R^{2p}} c_0(r_1')\cdots c_0(r_p')P_{X^{-1}(t,r_1'),\ldots,X^{-1}(t,r_p')}(r_1,\ldots,r_p)dr_1'\ldots dr_p' \tag{3.3}$$

where $P_{\xi_1,\ldots,\xi_p}$ is the probability density of a random vector $\xi_1,\ldots,\xi_p$, representing the probability distribution of Lagrangian trajectories in the fluid.

In oceanography, most of the work performed to date has focused on the first moment of c (i.e., on the mean concentration $\langle c\rangle$) and on the related probability density function for a single particle P_{ξ_1}. A few studies have considered the statistics of particle pairs (e.g., Bennett, 1984; Davis, 1985). Even in the simplest case of a single particle, though, the data are not sufficient to compute P_{ξ_1}, so that (3.3) cannot be used directly. Information on $\langle c\rangle$ can, in principle, be retrieved by combining the data with the equation for $\langle c\rangle$ obtained by averaging (3.1). The trouble with this approach is that the resulting equation for $\langle c\rangle$ involves terms such as $\langle \boldsymbol{u}\nabla c\rangle$; the equation for these terms in turn involves still higher order statistical terms, and so on in an unending hierarchy. This is the "closure" problem, one of the central problems in fluid dynamics. In practice, what is usually done is to "close" the equations for $\langle c\rangle$ at a chosen level using some kind of assumptions. The issue then becomes identifying the closed equations' appropriate form for the specific context under examination (e.g., see Molchanov and Piterbarg, 1992). As discussed in Chapter 1, the simplest form of closure is given by the advection and diffusion equation (1.4) where molecular diffusivity is replaced with turbulent ("eddy") diffusivity. An estimate of diffusivity can be obtained from the data, as a function of the velocity autocorrelation measured by buoys (e.g., Kraus and Boning, 1987). This form of closure is, strictly speaking, valid only if the flow is homogeneous in space and stationary in time, and if the time scales considered are longer than the time scales of the turbulence. Other more general and more widely valid equations have also been used in the literature. Examples are the elaborated form of the advection and diffusion equation proposed by Davis (1987) and stochastic models used to describe the motion of single particles (Thomson, 1986; Dutkiewicz et al., 1992).

One of the difficulties in using data from drifting buoys is that, whereas the data are inherently Lagrangian, the information oceanographers are interested in is often Eulerian (i.e., associated with a fixed point). Typically, oceanographers seek maps of simple statistics of the velocity, such as the mean flow and the variance, and of some turbulent transport quantities, such as the diffusivity. The knowledge of diffusivity as a function of space is of great importance for a number or reasons. First, it provides a direct picture of the nature of ocean turbulence, which is still not well understood (as discussed in Chapter 1). In particular, comparing diffusivity maps and maps of mean flow or velocity variance provides a way to test simple theories of turbulence, and eventually indicates how to improve them. Secondly, one must know diffusivity as a function of space, because it is an input of key importance for numerical models that simulate oceanic processes using equations (1.1)-(1.4) in Chapter 1.

The theoretical problem of determining Eulerian statistics from Lagrangian statistics is quite difficult, and it is still open (e.g., Li and Meroney, 1985; Babiano et al., 1985). Oceanographers take the simplest possible approach. They consider a set of measurements taken in a certain geographical region and assume that the region can be divided into smaller subregions (boxes) characterized by a space scale L, where the statistics are approximately homogeneous and stationary. All the data present in each box at all times can then be considered as representative of the same spatial point, and can be used to compute averages of the quantities of interest. In this way, the Eulerian statistics are computed from a combination of space and time averaging. The important question is, What happens when the hypotheses of homogeneity and stationarity inside the boxes are relaxed, as is expected to occur in a realistic situation? An extensive analysis regarding this problem has recently been done by Davis (1991b) in the context of the elaborated advection and diffusion equation. The following paragraph briefly summarizes some important points.

Stationarity can be relaxed fairly realistically provided the ocean is characterized by slowly varying fluctuations so that time averages, even though not constant, are representative of the particular ocean climate present during the measurements. Inhomogeneity could in principle be reduced inside each box by increasing the resolution, i.e., by decreasing L, the scale of the boxes. In practice, though, the uncertainty in the estimate of the statistical quantities also depends on L, so that a trade-off must be found between resolution and accuracy. The scale L must be large enough to give a reasonable uncertainty and small enough so that the statistical quantities computed in the box are meaningful.

It is important to note that biases can occur in estimating the statistical quantities as a consequence of both inhomogeneity in the sampling (array bias) and in the turbulent velocity (diffusion bias). This last type of bias reflects the observed tendency of drifting buoys deployed at a point to migrate toward regions of high turbulent energy. As shown by Davis (1991b), the size of these biases can be identified for mean velocity, but it appears to be much harder to identify for diffusivity. The use of other model equations for transport (or equivalently for particle motion) may help in identifying this bias or possibly suggest better estimators for the quantities of interest.

Finally, in some special cases the inhomogeneity of the statistical quantities can likely be solved explicitly. This can happen when general information is available on the spatial structure of the quantities, so that they can be approximated by space-functions dependent on a discrete number of parameters. An approach of this type has thus far only been applied to simple linear flows (e.g., Davis, 1985), but it is likely to also be useful for more complex flows, such as strong vortices or meandering currents, which play an important role in oceanography. The technique consists of estimating the parameters by using the data in conjunction with a model equation, such as some form of the advection and diffusion equation or a stochastic model for particle motion. The use of a stochastic model also provides a natural and straightforward way to filter the data.

PROSPECTIVE DIRECTIONS FOR RESEARCH

As is apparent from the preceding discussion, a number of key problems (e.g., the "closure" problem, determining Eulerian statistics from Lagrangian statistics, dealing with array bias and diffusion bias) are still open that relate to the use of Lagrangian data in the description of the ocean circulation. They suggest a variety of directions for statistical research, ranging from statistical analysis for oceanographic data to probabilistic modeling for processes in the ocean. Some specific considerations are the following:

1. Statistical methods for irregular and sparse observations, with emphasis on estimation of spectral and correlation characteristics (see Chapters 6 and 8);

2. Filtering and parameter estimation for random fields governed by randomly perturbed ordinary and partial differential equations, with emphasis on numerical methods for nonlinear filtering, spectral methods, and others;

3. The study of single-particle statistics in inhomogeneous and nonstationary turbulent flows;

4. The study of multiparticle statistics;

5. The Lagrangian approach to turbulence;

6. The derivation of closed-form equations for moments of passive scalars; and

7. The exploration of the time evolution of distributions of passive scalars, with emphasis on intermittence ("patchiness").

4
FEATURE IDENTIFICATION

A fundamental problem in oceanographic data analysis is the identification of features in image data: their shape, size, and motion. The data used in identification are typically satellite images, e.g., infrared or visible images from the NOAA polar-orbiting satellites or from synthetic aperture radar (SAR). Features are identified in order to quantify their statistics (e.g., ring size and frequency, front locations), to understand the evolution of the fields (e.g., ice leads and floes), and in successive images to infer motion in the field (e.g., sea surface temperature (SST)). Statistics of the features can be used to determine the accuracy of numerical models that describe the physics of the process. Feature identification can also be used to generate realistic fields from data with numerous gaps for assimilation into numerical models for prediction. Feature identification is usually complicated by the presence of instrument noise or geophysical (e.g., clouds) noise. Automation of feature identification using statistical measures is a primary issue; to date, few automated techniques have matched the success of a skilled analyst.

TRACKING OF FRONTS AND RINGS

The locations of major current systems and the location, tracks, diameters, and lifetimes of rings have been studied using infrared images from the Advanced Very High Resolution Radiometer (AVHRR) sensor on the NOAA polar-orbiting satellites. Brown et al. (1986) characterized the warm-core rings in the Gulf Stream system using 10 years of AVHRR data; a histogram of ring lifetimes showed two distinct peaks at 54 days and 229 days. Auer (1987) analyzed rings as well as the "north wall" of the Gulf Stream, defined subjectively as the location of the maximum SST gradient, using analysis charts derived from AVHRR images. Among other findings, Auer found that the position of the north wall had an annual signal, and that its interannual variability in position was comparable to its annual variability. Cornillon (1986) examined variations in the Gulf Stream position upstream and downstream of the New England Seamounts, again locating the north wall subjectively, and found that the meander envelope did not increase due to the seamounts, but that the mean path length did increase. Cornillon and Watts (1987) found that subjective identification of the north wall was more accurate than that enabled by any "conventional algorithm," such as the location of the maximum SST gradient, and found that the root-mean-square difference between the AVHRR-derived location and a traditional definition based on in situ temperature measurements was less than 15 km.

Ring motion is generally determined by the ring displacement over periods of tens of days, but there may be substantial changes in ring structure and motion over these time periods. Cornillon et al. (1989), in an attempt to determine the motion of warm-core rings relative to the motion of the Gulf Stream slope water, confined their analyses to pairs of observations separated by 36 hours or less. The ring outline was determined from AVHRR images, again by subjective methods, and the ring center was found by the best fit to an

ellipse. This fit to the ellipses was found to be better than both a center-of-mass estimate or the intersection of perpendicular bisectors from the ring edge. Absolute velocity estimates were derived from adjacent pairs of ring centers. The velocity of the slope water was determined by a subjective tracking of small SST features in pairs of images (horizontal velocity estimation is discussed in more detail below), and the difference between the velocity estimates was the desired result. The uncertainties in all of the motion estimates were quite large. A related problem is the determination of the ring characteristics and frequency of occurrence based on a series of line samples (as from a radar altimeter subtrack), where the spacing between tracks is as large as a ring diameter and the time between successive tracks is comparable to the time required to move to another track (an "aliasing" problem).

Mariano (1990) developed a method for combining different types of data to produce a map of a field that preserves typical feature shapes, rather than smearing them out as in an optimal estimate. Optimal estimates (generally known as "objective" maps in oceanography) minimize the expected squared error of the field value; Mariano's contour analysis produces instead an optimal estimate of the location of each contour of the field values. Thus it preserves the typical magnitudes of the field gradients; i.e., it preserves the shapes and sizes of rings and ocean fronts. Because the gradients affect the dynamics of the field in the simulation, the analyzed contour fields give more realistic input for assimilation into numerical simulation models. Mariano's method requires a pattern recognition algorithm to first delineate the contours in each type of data, before the optimal estimate of the final contour location can be made.

All of these statistical characterizations using images have in common the problem of detecting features in the presence of extensive *cloud contamination* or instrument noise; subjective methods have probably been most successful because the human eye can compensate for slight changes in the values of the field and locate a feature by its shape. The problem with subjective methods is that they tend to be labor intensive. A successful automated technique is highly desirable, especially for the case of analyzing large quantities of data (e.g., satellite observations or numerical model output). Ring studies have the additional problem of isolating an elliptically shaped feature that has numerous streamers and smaller eddies attached to it. The delineation of fronts is similar to a contouring problem: a single line must be designated in a noisy field, and the presence of closed contours must be determined to distinguish a ring from the front.

SEA ICE TRACKING

There are several problems in feature identification in sea ice for which good statistical estimators are needed. Some examples are given here. The motion of pack ice, using a feature-tracking method to determine velocities from a sequence of images, is similar to that of cloud motion or movement of water parcels (e.g., Ninnis et al., 1986). This problem is closely related to ocean velocity estimation, which is discussed below. Feature identification algorithms are needed to characterize ice floes (Banfield and Raftery, 1991;

see also Chapter 3 of NRC, 1991b) and leads (the open water between the ice floes): floe size distribution, and lead direction, spacing, and width distributions.

If one considers a set of markers on sea ice, their subsequent changes in position can be decomposed into four components: a translation, a rotation, an isotropic scaling, and a change in shape. An alternative decomposition would be into rigid body motion and deformation, and the deformation may be further decomposed into affine and nonaffine components. Shape statistics, concerned with the analysis of shapes such as these, includes the examination of a series of shapes evolving over time. In the context of polar oceanography, the emphasis is not so much on the shape itself—as it might be in biology where much of shape statistics originates—but rather on the motion and deformation of the shapes. The deformations and motions of various shapes must be reconciled with each other to establish the evolution of the entire field, and to infer something about the field dynamics.

A combination of feature identification and feature tracking is used to estimate the opening and closing of sea ice leads, which is necessary for models that estimate sea ice thickness (e.g., Fily and Rothrock, 1990). The object of this analysis is to produce an estimate of the fractional increase or decrease in size of sea ice leads from a pair of sequential SAR images. The first step in the estimation requires the designation of tie points between the same features in sequential images, which are determined by cross-correlations between subsets of the images. This procedure is quite similar to that required for estimation of ice motion. The next step requires the classification of the entire image into ice or lead, which is a statistical problem by itself, similar to that of flagging AVHRR images for cloud cover, or classifying AVHRR images by cloud type. The net increase or decrease in the area covered by the leads based on a comparison of the two classified images gives the required estimate.

ESTIMATION OF HORIZONTAL VELOCITIES FROM IMAGE SEQUENCES

Another oceanographic problem that might benefit from the application of advanced statistical methods is the estimation of horizontal ocean velocities using pairs of satellite images. One method of estimating these velocities is to track identifiable features in a tracer field, usually the sea surface temperature (SST; Emery et al., 1986). Other methods use the heat advection equation (1.4) (Kelly, 1989) or an assumption of geostrophic balance (Kouzai and Tsuchiya, 1990) to relate observed SST to the velocity field. SST images from the AVHRR have a horizontal resolution of approximately 1.1 km, with temporal separations of 4 to 8 hours. While clouds often obscure much of the ocean, there are occasionally periods of 1 to 3 days with relatively few clouds during which 4 to 12 images can be collected. Most of the velocity estimates assume that changes in SST are due to horizontal advection; however, other processes also change the SST seen by AVHRR: contamination by undetected clouds and fog, heating and cooling by the sun and air, vertical mixing and vertical motion, and changes in the top "skin" of the ocean (less than 1 mm thick). In the absence of these complications, the problem of estimating velocities would be one of mapping the location of all pixels in the first image onto the second image. It has been suggested that other statistical methods, such as simulated annealing (see, e.g., Chapter 2 in

NRC, 1992c), might produce such a mapping of individual pixels, but this has not been attempted to date.

The feature tracking method has been automated using a maximum cross-correlation (MCC) method, first applied by Emery et al. (1986) and derived from the methods used to track the motion of pack ice. The procedure is to cross-correlate a subregion of an initial image with the same-sized subregion in a subsequent image, searching for the location in the second image that gives the maximum cross-correlation coefficient. The size of the region searched in the second image depends on the maximum displacement that could be caused by reasonable velocities in the surface ocean. There is a trade-off between the spatial resolution of the velocity estimates and the statistical reliability of the cross-correlation. The small-scale features can be enhanced by the calculation of gradients or by high-pass filtering. It has been suggested that wavelet transforms might provide another way of first correlating larger-scale features and then smaller-scale features, but this has not been tried. Further references to the MCC method include Collins and Emery (1988), Kamachi (1989), Garcia and Robinson (1989), Tokmakian et al. (1990), and Emery et al. (1992).

Identifying features in consecutive images is not the most difficult problem in velocity estimation, although there is room for improvement here. Two related unresolved issues are ring motion (or rotation) and inferring velocity along isolines of the tracer field or in regions of small gradients. These flows produce only small changes in the tracer field, but the magnitudes of the velocities may be larger than those of the velocities that produce large changes in the tracer field. The MCC method can be modified to accommodate rotation of the features. Besides simply displacing the initial search region and calculating displacement, the initial region can be rotated through a reasonable range of angles (Kamachi, 1989; Tokmakian et al., 1990). However, the additional searches increase the chance of random high correlations, and the benefit is questionable. Emery et al. (1992) have investigated an alternate method of following rotation in closed rings and eddies, also noting that the basic method, without rotation, produces similar results.

Another method, which addresses the latter problem, solves the heat advection equation using inverse methods to find the velocity field most consistent with the change in SST fields observed in the two images (Kelly, 1989). The heat equation used, based on equation (1.4), is

$$T_t + uT_x + vT_y - m(x,y) = S(x,y), \tag{4.1}$$

where u, v are the horizontal velocity components, T_x, T_y are the horizontal derivatives of SST, T_t is the temporal derivative of SST, $S(x, y)$ is a term that describes SST fluctuations with relatively large spatial scales (which are not due to advection), and $m(x,y)$ is the misfit. As in the MCC method, there is an optimal temporal lag δ between images for the inversion: approximately 12 hours, compared to values of 4 to 6 hours preferred for the MCC method. Velocity fields that include the along-isoline velocity component can be obtained by adding constraints on the velocity solution, notably the minimization of horizontal divergence, with a weighting factor α relative to the heat equation (4.1), that is,

$$\alpha(u_x + v_y = 0). \tag{4.2}$$

Two-dimensional biharmonic splines were used as basis functions for the velocity fields in the inversion to give a continuous solution, unlike the feature-tracking methods, which give estimates at discrete grid points (Kelly and Strub, 1992). The spatial resolution of the solution depends on a parameter that sets the number of data per knot in the spline, and on the size of the subregion used to compute the SST gradients. A statistical challenge in this inverse problem is determining the best solution as a trade-off between the fit to the heat advection equation and the constraints. Although inverse theory methods exist to solve this problem more rigorously, it has not yet been done.

The horizontal velocity problem has been examined by many scientists and engineers. Other methods include the use of a single image in conjunction with the thermal wind equation, which relates horizontal SST gradients to vertical velocity shear (Kouzai and Tsuchiya, 1990). This method neglects salinity effects and requires an empirical relation between SST gradients and velocity from field data. Wahl and Simpson (1991) explored a variety of artificial intelligence methods for modifying the basic feature-tracking method and improving the cross-isoline solution. These methods have not been evaluated using field measurements.

The MCC and heat advection inverse methods have been compared by Kelly and Strub (1992) to in situ velocities from surface drifters and acoustic Doppler current profilers (ADCP), and to geostrophic velocities from the Geosat altimeter. They found that both methods produce velocity fields that captured the main features of the horizontal velocity field in a region of the coastal ocean approximately 500 km square. Both methods also underestimated the maximum velocities in the most energetic jets (velocities over 1 ms^{-1}). Detailed examination of the SST fields showed that in some cases the MCC method was not underestimating the displacements of identifiable features within the jet. Rather, drifters at 15-m depth within the jet were moving to locations beyond the SST feature in the second image. Thus, substantial errors in both methods occur because some of the largest velocities in the ocean do not produce observable SST changes. Although further modifications of these two methods or entirely new techniques might improve the estimates, these errors suggest that even a perfect mapping of SST fields would not give an accurate velocity field in regions with energetic jets. One promising approach is to incorporate independent velocity measurements into the estimate, either from radar altimeters or from drifters.

PROSPECTIVE DIRECTIONS FOR RESEARCH

Identifying features through the analysis of oceanographic data presents many opportunities for statistical research to contribute to progress on important physical oceanographic issues. The following particular issues exemplify some of the challenges for which statistical advances that improve on current approaches would be valuable:

1. Detection of SST fronts and rings (maximum gradients) in the presence of noise with a variety of spatial scales;

2. Characterization of rings or eddies by shape, frequency, and motion in a series of images or from a series of line samples, which may lead to aliasing of the feature motion;

3. Characterization of the evolution of ice floes and leads, using a time series of images. The emphasis is on inference of the dynamics of the field from the feature evolution and statistics; and

4. Estimation of oceanic velocity using a time series of tracer fields, where the relationship between the velocity field and the tracer is not unique and the velocity field is subject to some dynamical constraints.

5
VISUALIZATION

Scientific visualization has nearly become a cliché in recent years, as researchers apply increasingly sophisticated hardware and software tools to the task of data analysis. Techniques ranging from video animations of three-dimensional fields to simple two-dimensional line plots are often lumped under the term "visualization." In a sense, any visual representation of data may be considered visualization. However, a more useful definition would be more restrictive; visualization is the representation of data as a picture. This picture could consist of either static or evolving fields (animations).

The motivation for scientific visualization is the increasing availability and complexity of enormous observational data sets and numerical model output. Traditional line plots, tables of data, and other methods are inadequate to cope with the volume and complexity of these "data." Suitable visualization, by presenting the data as a picture, can allow the researcher to detect relationships and patterns much more quickly. This "illustrative" approach conveys information about relationships between components of the image simultaneously, rather than relying on a "discursive" or sequential approach using tables of numbers, sentences, and so on. The truism about a picture being worth a thousand words is applicable for many studies. In an effort to deduce the underlying processes responsible for the relationships between various physical phenomena, visualization tools will play an important role as scientists examine multidimensional data sets.

USES OF VISUALIZATION

The volume of data that can be collected by oceanographers has increased dramatically over the past 10 years. Although satellite sensors are the usual example, data rates from in situ instrumentation have also increased. For example, data storage technology now allows moorings to collect samples more frequently and for a longer time period. New instrumentation, such as spectroradiometers, are being deployed on moorings to measure upwards of 50 variables. Typical data sets now range from hundreds of megabytes to a few gigabytes or more.

Although the sheer volume of data may require visualization tools, an equally compelling need for improved visualization tools is the multitude of variables that are now being measured. Advances in ocean instrumentation have greatly increased the variety of processes that may be measured. For example, probes can now measure oxygen nearly continuously, rather than relying on bottle samples at a few discrete depths. High-resolution spectrometers measure phytoplankton fluorescence with much greater accuracy, resolving many pigments rather than just chlorophyll. The search for relationships becomes increasingly difficult as more data sets are added, and so analysis tools that simplify this process are essential. The need to examine complex relationships is not driven simply by our ability to measure numerous variables; rather, the importance of understanding the

interplay between biology, physics, and chemistry has driven the need for an interdisciplinary approach to data analysis.

Numerical models can now provide detailed three-dimensional views of the ocean. Such volumetric data are nearly impossible to analyze using traditional two-dimensional graphic techniques (see, e.g., Pool, 1992). The addition of the temporal dimension also requires animation tools to allow researchers to study model dynamics and evolution. Visualization tools play an important role in assessing model performance as well. For example, most model output has traditionally been discarded in an attempt to limit data volumes to manageable levels. However, specific events in model simulations often appear in just a few time steps, so that the ability to retain model output at every time step is useful for model diagnostics. The resulting large quantities of model output place a greater demand for sophisticated visualization techniques to search through the large volumes of data in an efficient manner that enables easy identification of the events.

CHALLENGES FOR VISUALIZATION

Visualization will continue to be important for oceanographic research as the ability to measure and model the ocean improves. Existing visualization tools, however, are inadequate for these tasks. Many deficiencies revolve around implementation problems and have been described in numerous NASA and other federal government reports (Botts, 1992; McCormick et al., 1987). For example, existing visualization packages are generally expensive and difficult to learn. Packages are usually not extensible, so that custom features cannot be added easily. Some tools cannot handle three-dimensional data sets or animations. One of the more difficult challenges is the ability to visualize evolving volumetric data, such as that produced by an ocean circulation model. It is very difficult to "see" into the interior of such volumes using present technology. Most commercially available packages that are designed for such volumetric data are capable of handling static images, such as automobiles. For many packages, visualizing three-dimensional systems that evolve over time is a difficult task. Such implementation deficiencies are slowly being addressed by the software vendors and developers.

The most troublesome aspect of existing visualization tools is that most of them break the link between the underlying data and the image on the screen. Although a researcher may be able to produce a sophisticated animation of the evolution of an ocean eddy, it is generally not easy to go from the animation on the computer screen back to the numbers that the various colors represent. As visualization is a tool to allow the detection of previously unknown relationships, it is still necessary to obtain quantitative information about the nature of the relationships. For example, if one notes a possible relationship between phytoplankton concentration and the strength of a density front in an eddy, it is desirable to examine the quantitative aspects of this relationship. Thus there must be techniques for excising subsets of the actual data for use in other analysis packages, such as statistical and plotting tools. Present visualization packages do not have probes or cursors that allow the user to examine the quantitative values of a three-dimensional image at specific locations,

nor do they have tools for graphically selecting subsets of visualized data (the equivalent of the "lasso" tool on the Macintosh).

Most earth science data are referenced to some system of Earth coordinates. As there is no standard way to carry such information along with the data, existing visualization packages either define their own format for such ancillary information or else discard it. It is vital that researchers be able to overlay different data sets on a geographic basis. A common example is the comparison of satellite maps of sea surface temperature and ship observations along a transect across the map. Again, most visualization tools do not retain this link to the underlying data. Visualization must include a link between the tools and an underlying database. This link must operate in both directions. That is, the visualization tool should be able to query databases to locate the raw data of interest for analysis, as well as maintain a database of the various visualization operations that were used to create a new, analyzed product. For example, an animation of vector winds and sea surface temperature might be created by querying a database. The steps used to create this animation would be stored along with the animation. Visualization tools can create large amounts of analyzed data that may be difficult to recreate without some type of audit-trail mechanism.

Currently, visualization tools are used largely in an exploratory manner, rather than for presentation to the research community. The high cost of color printing often prohibits the use of color imagery, and there is no established method for distribution of video animations. Occasionally, special sessions are held at scientific meetings for presentations of videos, but this approach reaches only a small fraction of the community. New methods for dissemination of visualizations must be established, as the existing print medium is not adequate. One approach would be to develop *animation* servers that are capable of storing and retrieving hundreds of video animations and other visualizations. For example, a research article might reference a video loop that is stored on the server, much as on-line library catalogs are stored now. With the planned increases in network capabilities, it would be possible to retrieve and view the animation on a local workstation. Such an animation could be an integral part of the paper and thus subject to peer review. If scientific visualization is made part of the publication process, it will no longer be just a tool for exploring data sets but a key component of scientific research and communication.

Lastly, color is often used in visualization to represent the underlying data. Most computer manufacturers have not invested in retaining color fidelity from device to device. For simple business graphics, variations in the shades of red from computer display to video tape to hard-copy printer may not be a serious concern. However, when this color represents specific data values in scientific applications, maintaining an exact shade of color across the breadth of output devices is essential for scientific research. This link to the data must also be maintained.

Visualization tools will likely increase in importance for oceanographic research as the volumes and complexity of data continue to increase. However, more attention must be paid to using these tools for their quantitative value, and not just for their ability to present complex relationships. This requires that these tools retain the links to the data that are used in the visualization process.

OUTSTANDING STATISTICAL ISSUES

One issue that could benefit from input from the field of statistics is the question of what method to use to interpolate irregularly spaced data to a regular grid in a manner that preserves the statistics of the field of interest (cf., NRC, 1991b). For example, satellite data generally consist of high-resolution data within measurement swaths, separated by hundreds or thousands of kilometers for which there are no data between swaths. Most interpolation methods smooth the data and minimize spatial gradients. It is desirable to retain as much of the full range of spatial scales as possible in the gridded fields.

Another issue that oceanographers are concerned with and that statisticians could contribute to is determining a method of identifying "interesting" events in the data that warrant a more detailed analysis. With small data sets, this can be accomplished by simply examining all of the data by various graphical techniques. For large satellite data sets or numerical model output, it is highly desirable to develop automated methods of locating such features. This can be done (with some success) for specific events with easily characterized features, but it is difficult when features are difficult to characterize concisely or do not possess simple characterizations.

6
INTERPOLATION, NONLINEAR SMOOTHING, FILTERING, AND PREDICTION

The topics of smoothing and filtering, commonly referred to as "data assimilation" in the oceanographic and meteorological literature, have attracted a great deal of attention of late. This emphasis on the combination of statistical with dynamical methods, relatively new to oceanography, arises as a natural consequence of the increasing sophistication of models, the rapid increase in available computing power, and the availability of new extensive data sets.

The most extensive of these newly available and soon to be available data sets are remotely sensed from space. Active and passive instruments operating in the microwave, infrared, and visible portions of the electromagnetic spectrum provide spatial and temporal coverage of the ocean unavailable from any other source, but present new challenges in interpretation. In particular, problems of filling in temporal and spatial gaps in the data, interpolating satellite data sets to model grids, and selecting a limited number of points from very large data sets in order to formulate tractable computational problems must be considered.

INTERPOLATION OF SATELLITE DATA SETS

Characteristics of Satellite Data

Different satellite instruments pose different problems, depending on spatial and temporal coverage, effects of clouds and rain cells, and viewing geometries. Characteristically, satellite data are sampled very rapidly (on the order of seconds or minutes). Data are acquired as areal averages along the satellite ground track, as in the case of the altimeter, which samples a region 10 km wide, or as areal averages of patches 5 to 50 km in diameter in swaths 1000 km wide in the cross-track direction, as in the case of the scatterometer or AVHRR. Spatially overlapping samples are taken on the order of 10 days later in the case of line samples, or on the order of 1 day later in the case of swaths.

The satellite altimeter, as indicated in Table 2.1 of Chapter 2, takes measurements roughly every 7 km along the track. Employing active microwave radar, the altimeter functions in both day and night hours, in the presence of clouds, or in clear weather. Two sets of satellite tracks, corresponding to ascending and descending orbits (i.e., orbits that cross the equator moving northward or southward, respectively) form a nonrectangular network that is oriented at an angle with respect to the parallels and meridians of latitude and longitude. The angles change as functions of the distance from the equator, as do the separations between adjacent tracks in the same direction (Figure 6.1). The irregular space-time sampling inherent in satellite measurements over an ocean basin raises important questions about aliasing and the range in wavenumber-frequency space that can be resolved

by the data. The problem is very difficult, and only a few attempts have thus far been made to address the issue (Wunsch, 1989; Schlax and Chelton, 1992).

Satellite instruments, such as AVHRR, that work in the visible and infrared range of the electromagnetic spectrum provide ocean observations only in the absence of clouds. Hence, maps based on these observations have gaps. One way of achieving full coverage of a specific ocean area is by creating composite images that combine data from different time periods (cf., NRC, 1992a). However, since the fields (for example, sea surface temperature) are time dependent, the composite images represent only some average picture of sea surface temperature distribution for the period covered. Therefore, it is important to know how this picture and its statistical properties are connected with the statistical properties of cloud fields, and how representative the composite image is with respect to the ensemble average of the temperature field (see, e.g., Chelton and Schlax, 1991).

Mapping Satellite Data: Motivation and Methods

For most applications, satellite data must be represented on a regular grid. The most common method of mapping satellite observations onto a geographic grid is by interpolating the data from nearby points at the satellite measurement locations. Given the complicated statistical geometry of oceanographic fields (see Chapter 2), such gridding may lead to considerable distortion. Therefore, it is important to study effects of intermittent and rare events, as well as effects of statistical anisotropy and inhomogeneity of oceanographic fields, on the gridding process.

Each interpolated value is typically computed from the *10 to 1000 closest data points*, selected from the millions of points typically found in satellite data sets. Common non-trivial methods of interpolating include natural or smoothing spline fits, successive corrections, statistical interpolation, and fitting analytical basis functions such as spherical harmonics. In all cases the interpolated values are linear functions of some judiciously chosen subset of the data.

Applications of natural splines and smoothing splines to interpolate irregularly spaced data are as common in oceanography as they are in most other fields of science and engineering. The methods have been well documented in the literature (e.g., Press et al., 1986; Silverman, 1985).

Successive corrections (Bratseth, 1986; Tripoli and Krishnamurti, 1975) is an iterative scheme, with one iteration per spatial and temporal scale starting with the larger ones. The interpolating weights are a function only of the scale and an associated quantity, the search radius (e.g., Gaussian of given width arbitrarily set to zero for distances greater than the search radius). This scheme is computationally very fast and adapts reasonably well to irregular data distributions, but does not usually provide a formal error estimate of the interpolated field, although it is straightforward to add one. Somewhat related is an iterative scheme that solves the differential equation for minimum curvature (Swain, 1976) of the interpolated surfaces with predetermined stiffness parameter, akin to cubic splines; however, the extension to three-dimensional data is not commonly available.

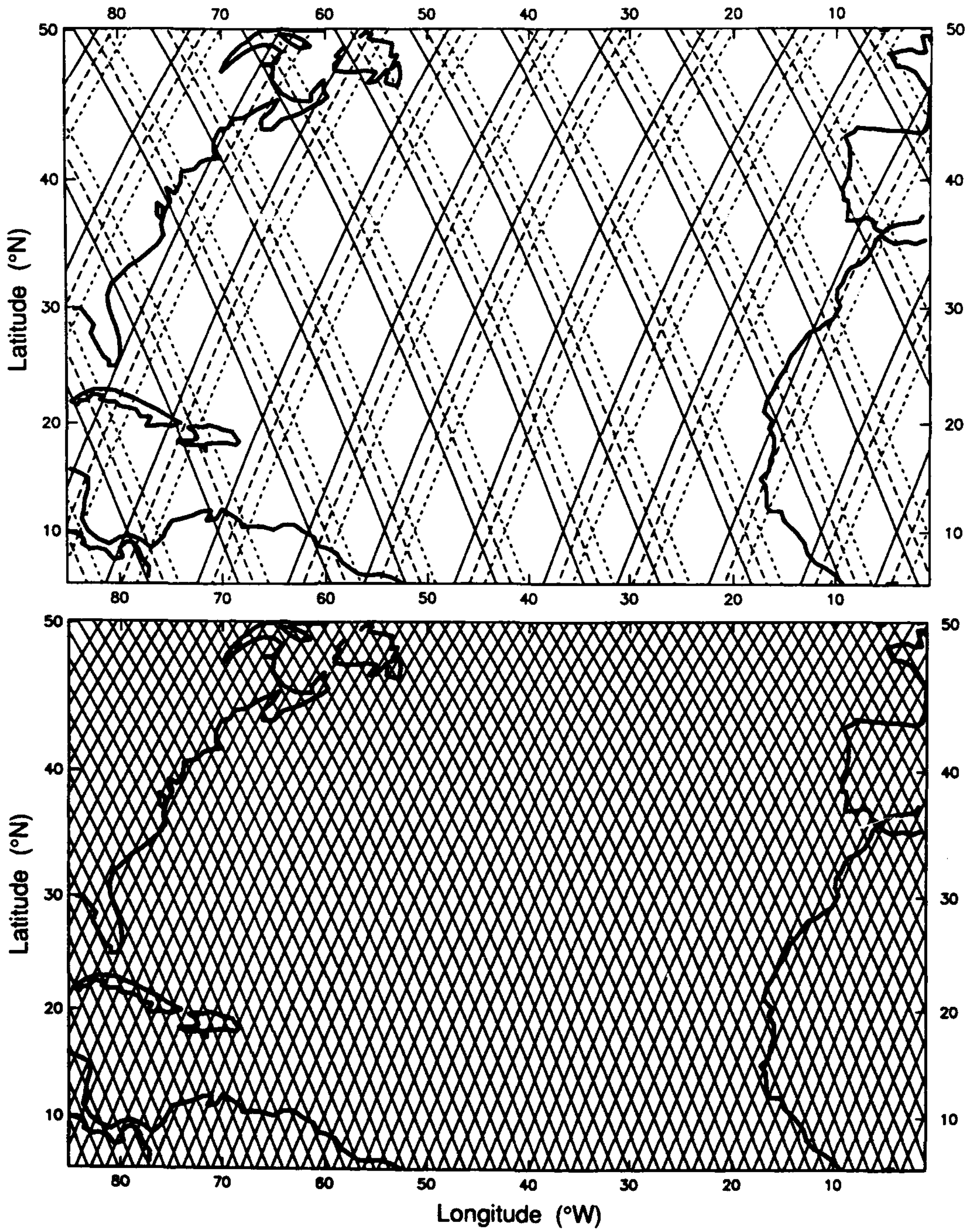

FIGURE 6.1 Example pattern of satellite ground tracks for the Geosat altimeter (see Douglas and Cheney, 1990, and Vol. 95, Nos. C3 and C10 of *J. Geophys. Res.*) with a 17-day exact repeat orbit configuration. Upper panel shows the ground tracks traced out during days 1 to 3 (solid lines), days 4 to 6 (dashed lines), and days 7 to 9 (dotted lines) of each 17-day repeat cycle. Note the eastward shift of a coarse-resolution ground track pattern at 3-day intervals. Lower panel shows the complete grid of ground tracks sampled during each 17-day cycle. SOURCE: Courtesy of Dudley Chelton, Oregon State University.

Statistical interpolation (Gandin, 1965; Alaka and Elvander, 1972; Bretherton et al., 1976), also referred to as optimal interpolation and most generally referred to as objective mapping (despite the fact that all of the techniques described here are objective), consists of least-squares fitting between interpolated and data fields. It assumes that estimates are known and available of the covariance matrix of the data with errors, and of the field to be interpolated. This is formally identical to ordinary least squares regression, in which the value of the interpolated field at a given point is assumed to be a linear function of the data at nearby points, and the moment and cross-product matrices are determined by assumptions about the spatial and temporal covariance of the underlying field. The formulas for the coefficients are derived simply by taking the expected values of the matrices in the ordinary least squares regression formula. Because matrix inversions are required for each set of estimates, the computational requirement is typically an order of magnitude larger than with the successive correction scheme. Formal error estimates are always given. Kriging (Journel and Huijbregts, 1978; NRC, 1992a) is a similar method in which the structure function rather than covariance is used to describe the data and desired field, with somewhat better adaptability to inhomogeneous statistics. The equivalence of objective analysis and spline interpolation was presented by McIntosh (1990).

Projecting the data on a space spanned by a convenient set of nonlocal basis functions is simple and well known, but there is no obvious choice of an efficient set of basis functions. The spherical harmonics commonly used for this purpose in meteorology do not form an efficient basis over the oceanic domain alone, requiring high-degree terms just to adapt to the domain. Recent efforts to define an equivalent set only over the oceans (Hwang, 1991) appear to have been successful.

The disadvantage of using L^2 norm minimizations is their relatively high computer resource requirement. An insidious consequence of this high resource demand is that in order to limit the problem to a size manageable with available computer resources, some researchers use too few data values or too small a region to achieve proper isolation of the length scales of signal and error. The disadvantage of schemes with fixed weights is clear: they are unable to adapt to data of varying accuracy, even though they do a decent job at adapting to inhomogeneous data distributions. The practical disadvantage for both objective mapping and successive corrections is that spatially inhomogeneous scales and anisotropy are not easily treated, and require breaking up the problem into several regional ones. This can lead to inconsistencies or other undesirable problems along the boundaries of adjacent regions. In the case of basis functions, most natural choices prove to be very inefficient in representing small-scale features; e.g., many higher-degree terms may be required to define a narrow jet such as the Gulf Stream.

DATA ASSIMILATION: USE OF DYNAMICAL MODELS FOR SMOOTHING AND FILTERING

As discussed in Chapter 1, it is not possible even with satellite data sets to provide complete initial and boundary conditions for the models in use today. This is partly due to physical considerations, such as the unknown details of air-sea exchanges, but the greatest

limitation on modeling studies today remains the sparsity of data, especially subsurface data that are inaccessible via satellites. It is therefore necessary to extract all available information from the data while, simultaneously, understanding the limitations on the applicability of any given data set.

Most data assimilation work to date has been based on least-squares formulations and the resulting linearized mathematical formalism. This can be justified rigorously for linear systems under fairly general conditions, assuming the initial error distributions are Gaussian. Within their realm of applicability, linearized methods have been quite successful. The ocean modeling community has a fair amount of experience with filtering and smoothing of linear models. The major remaining issues involve validation of statistical error models. These issues are most fruitfully considered in the contexts of specific problems. A review of data assimilation in oceanography can be found in Ghil and Malanotte-Rizzoli (1991). For a general overview, see NRC (1991a).

The use of ocean circulation models in smoothing and filtering of observational data has a relatively short history. Still, there have been a number of successful attempts (e.g., Thacker and Long, 1988; Gaspar and Wunsch, 1989; Miller and Cane, 1989). A recent excellent study is that of Fukumori et al. (1992). There has been, however, little systematic study of *nonlinear* smoothing and filtering in the context of ocean modeling. The ocean modeling literature naturally overlaps with the numerical weather prediction literature on this subject, and the two fields share a common interest in qualitative results, but systematic studies are few, and those that exist are elementary.

Direct approaches to applying statistically based data assimilation methods to nonlinear problems have so far been based on generalizations of linear methods. Variational methods used to date (e.g., Tziperman and Thacker, 1989; Bennett and Thorburn, 1992; Miller et al., 1992; Moore, 1991) have been derived from quadratic cost functions; i.e., the optimized solution is the one that minimizes some combination of covariances. This presupposes the notion that minimizing quadratic moments is the right thing to do in this context, even though the underlying distribution may not be unimodal. As one might expect, these methods work well in problems in which the nonlinearity is weak, or at least does not result in qualitatively nonlinear behavior such as bifurcation or chaos. Model studies have been performed on the Lorenz equations (Gauthier, 1992; Miller et al., 1992), which, for the most part, used covariance statistics and linearized methods. (An application in which third and fourth moments were calculated explicitly was presented by Miller et al. (1993), but it is unlikely that this method has any wider applicability). Gauthier (1992) and Miller et al. (1993) discuss in detail the pitfalls in filtering and smoothing of highly nonlinear problems. In those cases, the implementation of variational methods results in extreme computational difficulty.

The solution to the nonlinear filtering problem for randomly perturbed dynamical systems is well understood theoretically (see Rozovskii, 1990). It can be reduced to a solution of the so-called Zakai equation, a second-order stochastic parabolic equation. It describes the evolution of the non-normalized density of the state vector conditioned upon observations. Smoothing and prediction are technically based on the Zakai equation and the so-called backward filtering equation (see Rozovskii, 1990). In the last decade substantial progress has also been made in numerical studies of the Zakai equation (see, e.g.,

Florchinger and LeGland, 1990). However, this theoretically perfect approach has some practical limitations. In particular, the dimension of the spatial variable for the Zakai equation is equal to the dimension of the state vector. This is clearly impractical for modern dynamical ocean models that have thousands, if not hundreds of thousands, of state variables.

It appears that the most promising approach to this problem is development of hierarchical methods that would involve Kalman-type filtering where possible and refinement of the first-level coarse filtering by application of intrinsically nonlinear procedures when necessary. These require further research on numerical approximation for Zakai-type stochastic partial differential equations, including development of stochastic versions for multigrid methods, wavelets, and so on.

While true nonlinear filtering will not find direct application to practical ocean models in the near future, guidance from solutions of simplified problems can be expected. Further, there may be approximations to the Zakai equation in terms of parametric representations to solutions that are more versatile than those derived from methods explored earlier.

Overall, it appears that numerical methods for stochastic systems are developing into an exciting area of science that is of importance to oceanographic data assimilation.

INVERSE METHODS

Some oceanographers consider that, in some larger sense, all of physical oceanography can be described in terms of an inverse problem: given data, describe the ocean from which the data were sampled. Obviously direct inversion of the sampling process is impossible, but the smoothing process is occasionally viewed as some generalized inverse of the sampling process, with the laws of ocean physics used as constraints (see, e.g., Wunsch, 1978, 1988; Bennett, 1992).

It has become common in oceanography and dynamic meteorology to solve the smoothing problem by assuming that the system in question is governed exactly by a given dynamical model. Since the output of many dynamical models is determined uniquely by the initial condition, the problem becomes one of finding the initial conditions that result in model output that is closest to the observed data in some sense; the metric most commonly used has been least squares. These problems are usually solved by a conjugate gradient method, and the gradient of the mean square data error with respect to the initial values can be calculated conveniently by solving an adjoint equation. For that reason, this procedure is often referred to as the adjoint method. (see, e.g., Tziperman and Thacker, 1989). This is formally an inverse problem, i.e.: when given the outputs in the form of the data, find the inputs in the form of the initial conditions.

There are many significant problems in physical oceanography that bear specific resemblance to what is formally called inverse theory in other fields such as geophysics. These include estimation of empirical parameters (e.g., diffusion coefficients) and the design of sampling arrays to yield the most detailed picture of the property being sampled. Problems such as these, along with others that fall within the strict category of smoothing and filtering, are described in detail in the volume by Bennett (1992).

PROSPECTIVE DIRECTIONS FOR RESEARCH

There are many opportunities for statistical and probabilistic research regarding interpolation, smoothing, filtering, and prediction associated with oceanographic data. The following are some of the contexts that present challenges:

1. Filtering and smoothing for the systems in which the dynamics are given by discontinuous functions of the state variables;

2. Parameter estimation for randomly perturbed equations of physical oceanography;

3. Alternative numerical and analytical approaches to the least-squares approach for nonlinear systems;

4. Hierarchical methods of filtering, prediction, and smoothing;

5. Spectral methods for nonlinear filtering (separation of observations and parameters);

6. Multigrid and decomposition of the domain for Zakai's equation; and

7. Application of inverse methods for (a) data interpolation, (b) estimation of empirical and/or phenomenological parameters, and (c) design of sampling arrays.

In particular, progress in answering the following questions would certainly be beneficial:

1. What is the best way to solve the smoothing problem in cases where the dynamics are given by discontinuous functions of the state variable? Such examples are common in models of the upper ocean in which convection takes place. Possibly the best ocean model known, that of Bryan (1969) and Cox (1984), deals with this problem by assuming that the heat conductivity becomes infinite if the temperature at a given level is colder than it is below that level. The result is instantaneous mixing of the water, to simulate the rapid time scale of convection in nature. This can be viewed as an inequality constraint on the state vector; i.e., some regions of state space are deemed to be inadmissible solutions of the problem. Such problems are treated in the control theory literature (see, e.g., Bryson and Ho, 1975), but the engineering methods are not conveniently applicable to high-dimensional state spaces.

2. If the least-squares approach is inadequate for highly nonlinear systems, what would be better?

3. What is the best way to apply solutions of the nonlinear filtering problem to more complex systems? Might it be possible to implement the extended Kalman filter for a

relatively simple system and use the resulting covariance statistics in a suboptimal data assimilation scheme for a more detailed model? In general, how might the hierarchical approach suggested in the section above on data assimilation (also cf., NRC 1992a) be implemented?

4. When should one statistical method be applied as opposed to another? What diagnostics are there to help make decisions on suitable methods? Answers to such questions could be compiled in a handbook on statistical analysis of oceanographic and atmospheric data, could include such things as definitions and methods of statistical parameter estimation, and could discuss such questions as, e.g., What do these parameters convey?

5. What statistical methods can be used for cross-validating data that take inherent averaging errors into account, and that provide estimates of their magnitude? With the advent of remote sensing, data comparison (Chapter 7) is not limited merely to measurements and model verification, but involves cross-validation of different sensors or assimilation of data into models for quality assessment (see NRC, 1991a). In such analyses, each data set contains errors that are inherent to the averaging process. As Dickey (1991, p. 410) has noted:

> One of the major challenges from both the atmospheric and ocean sciences is to merge and integrate in situ and remotely sensed interdisciplinary data sets which have differing spatial and temporal resolution and encompass differing scale ranges . . . . Interdisciplinary data assimilation models, which require subgrid parametrizations based on higher resolution data, will need to utilize these data sets for applications such as predicting trends in the global climate.

7
MODEL AND DATA COMPARISONS

Oceanographers often have available multiple independent estimates of the various geophysical quantities of interest (e.g., sea surface temperature, surface winds, surface humidity, sea level, velocity, etc.). The sources of such estimates might be in situ observations, satellite-based observations, numerical model simulations, or so-called analyzed fields. The latter may consist of regularly gridded estimates constructed by subjective (i.e., hand-drawn) or objective (i.e., computer generated by some objectively prescribed interpolation algorithm) analysis of irregularly spaced observations. Alternatively, analyzed fields may be constructed from a numerical model forecast, adjusted to be consistent in some least-squares sense with all available observations acquired since the previous "analysis time." Independent estimates of the same quantity are never precisely the same, and small differences can sometimes have a profound influence on the scientific interpretation or application of the geophysical field. An important statistical problem in oceanography is therefore development of techniques for quantitatively evaluating the degree of similarity or difference between independent estimates of a multidimensional field. This includes cross-comparisons between different observational data sets (e.g., in situ vs. satellite), comparisons of model simulations with observations, and comparisons between different model simulations.

An example of a geophysical quantity that illustrates the kind of problems that can be encountered in comparisons of different observational data sets is sea surface temperature (SST). Temporal variations of SST are generally dominated by the seasonal cycle, which may have an annual range of 5° to 10° C or more at any particular geographical location. Interannual deviations from the local seasonal cycle typically have magnitudes of only about 0.5° C. Such small anomalies in SST can have a significant effect on climate. Even the El Niño phenomenon that affects weather patterns on a global scale can be initiated by an SST anomaly in the eastern tropical Pacific of only a degree or two. It is very difficult to estimate SST to an accuracy of 0.5° C by any of the means currently available. Since the actual SST is not known on ocean-basin scales, it is difficult to assess the accuracy of the several different estimates available. Attempts to determine the accuracy of satellite estimates of the SST field are often made by comparisons with in situ observations from ships and buoys or with other satellite-based estimates (e.g., Bernstein and Chelton, 1985). In the case of in situ observations, comparisons are complicated by the sample size and distribution. The data are not uniformly distributed geographically or temporally. Observations tend to be concentrated along standard shipping routes and are generally more sparse during severe wintertime weather conditions. Moreover, in situ observations can differ from satellite estimates because of measurement errors and because of smaller-scale variations that are spatially averaged in satellite measurements. Comparisons between two different satellite estimates of SST are complicated by a common source of error, atmospheric effects on the radiance emitted from the sea surface, which obscures the errors in both data sets.

Systematic errors, particularly in satellite data, create biases in the simplest statistical measures, be they spatial or temporal averages. In addition to the problem of limited sample size discussed above (see also Preisendorfer and Barnett, 1983), such gross statistics can obscure important characteristics of the differences such as geographical or temporal biases (see, e.g., Barnett and Jones, 1992). For the SST example above, such biases may arise from systematic errors in the algorithms applied to correct for atmospheric effects on satellite estimates of SST. As an example, volcanic aerosols injected into the atmosphere by the El Chichon volcano in 1982 contaminated infrared-based satellite estimates of SST within about 30° of the equator for a period of about 9 months. As another example, microwave-based satellite estimates of SST have been found to be biased upward in regions of high surface winds because of incomplete corrections for the effects of wind speed on ocean surface emissivity.

Evaluation of numerical model simulations, either through comparisons with observations or by comparisons with other model simulations, presents additional problems. Models produce a large number of output variables on a dense space-time grid. An ocean circulation model, for example, typically outputs current velocities, temperatures, and salinities at a number of different depths, as well as the sea surface elevation. It is not reasonable to expect present models to reproduce the details of the actual circulation, but one hopes that basic statistics such as the mean or variance of some characteristics of the actual circulation are well represented by the model. Assessing the strengths and weaknesses of a model is thus complicated by the large number of possible variables that can be considered. For example, present global ocean circulation models can reproduce the statistics of sea level variability with some accuracy but generally underestimate the surface eddy kinetic energy computed from surface velocities (e.g., see Morrow et al., 1992). A model that successfully represents the statistics of some geophysical quantity at one level may misrepresent the statistics of the same quantity at a different level. An even more stringent assessment of the performance of a model is how accurately it represents cross-covariances between different variables (which can be shown to be related to eddy fluxes of quantities such as heat, salt, or momentum). Some of these issues are discussed by Semtner and Chervin (1992) with regard to comparisons of numerical model output to satellite altimeter estimates of sea level variance and eddy kinetic energy. The overall goal of such comparisons is to guide further research in an effort to develop more accurate numerical models.

The types of questions that need to be addressed by techniques for comparing two different geophysical fields, whether they consist of observations or model simulations, are indicated by the following:

1. How, where, and when do the two independent estimates of a field differ?

2. Are the differences statistically significant? Addressing this question may lead to development of appropriate bootstrap techniques for estimating probability distributions.

3. What statistical comparisons are most appropriate for evaluating a model?

8
NON-GAUSSIAN RANDOM FIELDS

For purposes of statistical analyses, oceanographic fields are usually assumed to be Gaussian, stationary, and spatially homogeneous, and their statistical description is limited to the calculation of wavenumber spectra. However, since oceanographic stochastic partial differential equations (see Chapter 2) are nonlinear or bilinear, the statistics of the fields depart from such simple models. The nonlinearity is due mainly to advective terms such as $(\mathbf{u} \cdot \nabla)\mathbf{u}$ where $\mathbf{u}$ is the velocity vector for water motion. In some cases, specifically for surface gravity waves, the nonlinear nature of the fluid motion is due to nonlinear boundary conditions: water motion is described by a function and is governed by the Laplace equation, while the (kinematic) boundary condition expressing the continuity of the free surface is nonlinear. As a result, closed equations for various statistical moments of the fields cannot be rigorously derived. Pertinent definitions and statistical problems are reviewed in two comprehensive volumes on statistical fluid dynamics by Monin and Yaglom (1971, 1975). A review of statistical geometry and kinematics of turbulent flows is given by Corrsin (1975). Walsh (1986) and Rozovskii (1990) provide introductions to stochastic partial differential equations.

One of the most important and least understood features of oceanographic processes is the intermittent (rare) occurrence of special or catastrophic events. These include (in order of increasing scale) appearance of white caps at the crests of exceedingly steep and breaking surface gravity waves, patches of small-scale turbulence left by breaking internal waves, the shedding of mesoscale rings and eddies by large-scale currents (such as the Gulf Stream or the Agulhas current), and the occurrence of localized anomalies in SST including El Niño events with a time interval on the order of years. Such events play a very important role in the overall dissipation of kinetic energy, and in the transport of heat, salt, and other quantities by ocean currents, as well as in the exchange of energy, momentum, and chemical quantities across the air-sea interface. In terms of the primitive equations describing individual realizations of oceanographic fields, such events may often be viewed as singularities developing in the process of a field's evolution. Statistical analysis and modeling of such events are highly desirable. The use of quantile estimates might be investigated, especially for information in the tail of the distribution. The statistical geometry of these intermittent events is poorly understood, and improved understanding can be achieved by accounting more fully for the non-Gaussian nature of oceanographic fields.

Considerable progress in statistical modeling of geophysical "turbulent" fields has been achieved using ideas of multifractal processes (e.g., Schmitt et al., 1992). However, most of this work is related to atmospheric phenomena (Lovejoy and Schertzer, 1986; Schertzer and Lovejoy, 1987). A review of various problems arising in remote sensing, geophysical fluid dynamics, solid earth geophysics, and ocean, atmosphere, and climate studies can be found in Schertzer and Lovejoy (1991).

The special case of weak turbulence (when the nonlinear terms are of second order with respect to the linear terms in the governing equations) deserves particular attention, for it is encountered in many oceanographic problems and can be treated by small-perturbation

techniques. Examples of weak turbulence include two-dimensional and geostrophic turbulence and surface gravity waves. Weak turbulence theory in its present form (Zakharov et al., 1992) permits derivation of kinetic equations describing energy exchanges (and exchanges of other quantities) among Fourier components, as well as derivation of higher-order spectra (bispectra, etc.) representing Fourier transforms of various statistical moments. Initially, this theory was developed for surface gravity and capillary waves (Hasselmann, 1962; Zakharov, 1984). However, statistical phenomena in waves (e.g., the existence of Kolmogorov-type spectra, the intermittency of breaking waves, and so on) have analogies in other oceanographic fields. The elegant Hamiltonian formulation of nonlinear wave dynamics (Zakharov, 1984; Zakharov et al., 1992) is a powerful tool for studies of fundamental statistical properties of turbulent fields.

To better characterize the scope of statistical issues that the weak turbulence theory or alternative statistical approaches could address, a brief review of some issues related to wind-generated surface gravity waves is in order. Until recently, statistical studies of field geometry were dominated by the work on Gaussian fields. Longuet-Higgins (1957, 1962, 1984) studied a large variety of geometrical properties of such fields with application to sea surface waves. Among other problems, he considered statistics of specular points (the points at which the gradient of the field is either zero or is specified depending on a viewing angle) and of the wave envelope, which play an important role in wave dynamics and analysis of sun glitter and radar backscatter from a wind-disturbed sea surface. A rigorous mathematical analysis of envelope statistics, high-level excursions, field maxima, and other geometrical properties of random two- and multi-dimensional Gaussian fields is presented by Adler (1981). Some of these results have been successfully employed in sea wave studies. Specifically, the theory of level crossings by two- and three-dimensional Gaussian- and Rayleigh-distributed fields was employed to estimate statistics of whitecaps (breaking waves) and of wave trains (Glazman, 1986; Glazman and Weichman, 1989; Glazman, 1991). Observations indicate that whitecaps occur in clusters. Hence, the use of a simple Poisson distribution (Glazman, 1991) for whitecap occurrence, which is known from the theory of high-level excursions by the (Gaussian) wave slope field, may be insufficient. The statistical theory of cluster point processes may be of great help here.

Linear methods are intrinsic for Gaussian stationary processes, and Fourier analysis is a natural tool to use in the resolution of stationary random fields. These yield a global resolution. However, in many situations, a resolution that is better adapted to local behavior would be more appropriate and interesting. This could be local behavior in time or local spatial behavior. One attempt in this direction makes use of wavelet transforms, which are in effect local filters of the field (Farge, 1992). Such a method amounts to a linear analysis of the field, although it could presumably be adapted to types of nonlinearity.

In the last few years, significant research effort in probability and statistics has been directed toward the development of models of non-Gaussian and time-varying random fields. Examples include stable fields; functionals of Gaussian, stable, and other fields represented via multiple integrals; density processes and measure-valued diffusions; and fields described by nonlinear stochastic differential equations. Applications of this research to oceanographic phenomena would be of interest to oceanographers since the fields they study are frequently non-Gaussian and time-varying random fields.

One of the questions that arises in ocean remote sensing concerns the probability density function (pdf) for the heights of specular points and for the slopes and curvature radii of the surface. These pdfs are essentially non-Gaussian. A particularly interesting problem is statistically characterizing the asymmetry of the sea surface shape about the horizontal plane coincident with the mean sea level. This asymmetry is responsible for the deviation of the mean height of the specular points from the mean (zero-valued) height of the surface itself. As a result, an error bias (known as the sea-state bias) appears in altimeter measurements of the sea level. Mathematical analysis of such non-Gaussian surface properties is based on approximate joint pdfs for surface height and slopes. Following the work by Longuet-Higgins (1963) in which a truncated Gram-Charlier series expansion for the joint pdf was derived, the sea-state bias has been related to various spectral moments (Jackson, 1979; Srokosz, 1986) and ultimately expressed in terms of wind-wave generation conditions. While a simplified case of a one-dimensional surface has been studied, a two-dimensional case needs additional effort. The estimation of joint pdfs for dependent random sequences is reviewed, e.g., by Rosenblatt (1991). Further statistical effort in this direction could greatly facilitate analysis of biological and other oceanographic multidimensional processes.

The arrival of supercomputers opens new avenues for numerical modeling of complex processes. Now, for instance, numerical simulation of electromagnetic scattering by individual realizations of the random sea surface has become feasible. In this regard, simulated non-Gaussian random fields that satisfy basic conservation laws of fluid dynamics represent a great interest. A possible way of constructing individual realizations of a random field might be via the use of Wiener-Hermite polynomials (i.e., the Wiener-Ito expansion (Major, 1980)) in which the functional coefficients are determined on the requirement that the field yields the correct cumulants up to a certain order. Although bispectra (in the frequency domain) for surface gravity waves have been known since the work by Hasselmann et al. (1963), cumulants above second order for the surface's spatial variation have not been studied. In the literature on large-scale ocean dynamics (two-dimensional and geostrophic turbulence), the Wiener-Ito expansion has never been used, although it appears to be most relevant. Estimation of the cumulant spectra is discussed in the pioneering work of Brillinger and Rosenblatt (1967). See also Rosenblatt (1985) and more recent material in Lii and Rosenblatt (1990).

STATISTICAL RESEARCH OPPORTUNITIES

There are many statistical research opportunities in the realm of non-Gaussian physical oceanographic random fields on which progress would be desirable. Some specific topics worthy of investigation are the following (also see related issues in Chapter 2):

1. Models of non-Gaussian and time-varying random fields: (a) probabilistic analysis of different models of non-Gaussian or nonstationary or time-varying random processes and fields (e.g., stable fields, measure-valued diffusions, density

processes, non-Gaussian generalized fields, and so on), (b) structure of random fields with long-range dependence, and (c) non-Gaussian time series;

2. Theoretical models and techniques of simulation of non-Gaussian random fields with prescribed statistical properties, for example, (a) known moments up to some order, (b) known tail behavior of multivariate probability density functions, and (c) known statistics of extremes;

3. Extrema, sample path behavior, and geometry for non-Gaussian random processes and fields;

4. Inference and analysis of point processes with applications to oceanographic data;

5. Analysis of the Navier-Stokes system driven by Gaussian and non-Gaussian white noise;

6. Analysis of random fields that appear as solutions of stochastic partial differential equations (of special interest are equations driven by non-Gaussian noise or noises over a product of time-space and location-space);

7. Wavelet analysis of random fields with application to oceanographic problems; and

8. Statistical problems for non-Gaussian data (see models of particular interest in 2. above): (a) modeling (model identification, parameter estimation, and so on), (b) data analysis of irregularly sampled points on a field, (c) quantile estimation from dependent stationary processes and fields, (d) estimation problems for random fields given the types of sampling or observational layouts that are typical in oceanography, and (e) estimation problems for samples from non-Gaussian random fields.

9
ENCOURAGING COLLABORATION BETWEEN STATISTICIANS AND OCEANOGRAPHERS

Offered for the purpose of encouraging successful collaborations between statisticians and oceanographers, the following conclusions, observations, and suggestions are based on information that the Panel on Statistics and Oceanography gathered in this study, on the panel discussions that took place in preparing this report, and on the panelists' own experience and knowledge concerning cross-disciplinary research and collaborative efforts. The panel believes understanding and appreciating these matters are as important to the encouragement and accomplishment of statistical research in physical oceanography as are the descriptions of statistical research opportunities discussed in Chapters 2 through 8.

CONCLUSIONS

1. There are many opportunities for statistical research in biological, chemical, geological, and physical oceanography, far more than this report can address (owing to constraints of time and resources). This report thus represents a first step, focusing on challenging statistical issues in physical oceanography. However, the statistical problems it describes are universal, and progress on them would benefit the other oceanographic disciplines and also contribute to a better understanding of the coupled ocean-atmosphere system, weather patterns, and global climate change.

2. Many sophisticated statistical techniques are used routinely in physical oceanography. Nevertheless, in numerous general areas collaboration between oceanographers and statisticians could contribute to improving currently used models, analysis techniques, data assimilation methods, visualization methods, and so on. Examples of such areas identified in this report include multiple-scale variability of oceanographic fields; use of Lagrangian data in descriptions of ocean circulation; ocean feature identification; pictorial representation of oceanographic data; interpolation, smoothing, filtering, and prediction in the context of oceanographic data; comparison of oceanographic models and data; and non-Gaussian, nonstationary random fields.

3. Identifying research areas of mutual interest and need is basic to achieving results of genuine value to all participants in cross-disciplinary projects; another crucial requirement is providing an environment that encourages and sustains individuals who embark on collaborative research. Although exploring this second issue was beyond the scope of this study, the panel became increasingly aware during its deliberations of just how difficult it can be to engage in truly collaborative, cross-disciplinary work. There are many possible reasons for such difficulties (see, e.g., NRC 1990a): different parties in a cross-disciplinary collaboration may have different motivations or different disciplinary imperatives; there may be institutional impediments due to the traditional organization of separate disciplines within

an institution; there may be inherent obstructions to peer-reviewed funding or publishing of cross-disciplinary research (for instance, in defining what constitutes a peer); and there may be contextual scientific obstacles (since the multifaceted system under study may not fit into traditional categories for scientific investigation).

Without attempting to specify particular remedies, the panel includes below a few generic observations and outlines some possible initial approaches to encouraging collaborative research, especially between statisticians and oceanographers. The recent publication of several excellent studies and reports addressing cross-disciplinary research in various contexts (e.g., NRC, 1987; Institute of Mathematical Statistics, 1988; NRC, 1990a; see also Goel et al., 1990; Gnanadesikan, 1990; Hoadley and Kettenring, 1990), together with heartening signs of an improving environment for such activities (Crank, 1993; Harris, 1993), suggests that attention to the value of collaborative research is increasing and that work toward facilitating it will be ongoing.

OBSERVATIONS AND SUGGESTIONS

1. The need for clear communication and substantive interaction among collaborating researchers from different disciplines suggests the desirability of their working together at the same physical location for a significant period of time on specific problems to which both parties can contribute needed expertise. Funding agencies and research institutions could stimulate such interactions (a) by sponsoring workshops on well-delineated topics — drawn, for example, from the research areas discussed in this report — that are best addressed by a collaborative effort; (b) by providing for postdoctoral fellowships, senior research sabbaticals, and graduate student residencies that would enable statisticians to work with oceanographers at oceanographic research institutions; and (c) by sponsoring a series of one- or two-week short courses on oceanography for statisticians in which specialists would review selected topics and indicate open areas of research. It is much more likely that statistical research on one of the physical oceanographic challenges described in this report will produce valuable results if that research involves continuous interaction with an oceanographer who is versed both in the nuances of that challenge and in the practical oceanographic realities surrounding it.

In all such considerations, the panel encourages active cooperation between statisticians and oceanographers at agencies that fund research in these disciplines.

2. Effectively communicating the results of successful collaborative research — and thereby increasing understanding of its value in addressing complex problems — includes having the results published in journals that are well regarded in the relevant disciplines. The panel suggests that, as an initial step, one or more of the major statistical journals could publish a special section or issue on statistics and oceanography designed to increase awareness of the research opportunities in that area. This would encourage interaction between statisticians and physical oceanographers, increase the visibility of the results of successful collaboration, and set a precedent that could stimulate other highly regarded disciplinary journals to publish statistics and oceanography cross-disciplinary papers.

3. Promoting and nurturing cross-disciplinary research in statistics and physical oceanography, which will likely involve broadening the educational base of prospective researchers as well as the criteria by which their later efforts are rewarded, can be fostered now (a) by university statistics departments that stimulate cross-disciplinary interactions and learning and encourage statistics undergraduate and graduate students to obtain an "applied" minor in some other area, with oceanography being but one possibility (others being physics, engineering, geology, and so on), and (b) by funding agencies that promote a broader orientation in graduate and undergraduate statistics education.

It is likely that many people will be encouraged to undertake the significant efforts interdisciplinary statistics and oceanography research requires if funding agencies offer prospective cross-disciplinary collaborators some likelihood of obtaining research support, if recognized journals in an individual's discipline offer sufficient flexibility in publishing such cross-disciplinary research papers, and if research institutions accord cross-disciplinary research the same level of professional recognition (in promotion and tenure considerations) as is currently given to research in the individual disciplines.

Many major national and global concerns involve scientific research challenges that are cross-disciplinary in nature, with weather prediction and global climate change being but two examples related to the focus of this report. Encouraging the pursuit of such cross-disciplinary research opportunities can benefit both science and society by focusing scientific attention on research issues relevant to societal concerns. Encouraging the pursuit of cross-disciplinary research opportunities in statistics and oceanography will certainly benefit both disciplines: application of sophisticated statistics techniques will lead to better descriptions and improved dynamical understanding of oceanographic phenomena, and the statistics research challenges presented by oceanographic issues will inspire the development of new statistical techniques.

BIBLIOGRAPHY

Adler, R. J. 1981. *The Geometry of Random Fields*. Books on Demand, Ann Arbor. 292 pp.

Alaka, M. A., and R. C. Elvander. 1972. Optimum interpolation from observations of mixed quality. *Mon. Wea. Rev.* **100**:612-624.

Arvidson, R., F. Billingsley, R. Chase, P. Chavez, M. Devirian, J. Estes, G. Hunolt, J. C. Klose, G. Ludwig, F. Mosher, and W. Rossow. 1986. *Report of the EOS Data Panel.* NASA Tech. Memo 87777. National Aeronautics and Space Administration, Greenbelt, Md.

Auer, S. J. 1987. Five-year climatological survey of the Gulf Stream system and its associated rings. *J. Geophys. Res.* **92**:11709-11726.

Babiano, A., C. Basdevant, and R. Sadourny. 1985. Structure functions and dispersion laws in two-dimensional turbulence. *J. Atm. Sci.* **42**:941-949.

Baggeroer. A., and W. Munk. 1992. The Heard Island feasibility test. *Phys. Today* **45**(9):22-30.

Banfield, J. D., and A. E. Raftery. 1991. Identifying ice floes in satellite images. *Nav. Res. Rev.* **2**:2-11.

Barnett, T. P., and K. Hasselmann. 1979. Techniques of linear prediction, with application to oceanic and atmospheric fields in the tropical Pacific. *Rev. Geophys. Space Phys.* **17**:949-968.

Barnett, T. P., and P. D. Jones. 1992. Intercomparison of two different southern hemisphere sea level pressure datasets. *J. Clim.* **5**:93-99.

Bennett, A. F. 1984. Relative dispersion: local and nonlocal dynamics. *J. Atm. Sci.* **41**:1881-1886.

Bennett, A. F. 1992. *Inverse Methods in Physical Oceanography*. Cambridge University Press, Cambridge.

Bennett, A. F., and M. A. Thorburn. 1992. The generalized inverse of a non-linear quasigeostrophic ocean circulation model. *J. Phys. Oceanogr.* **22**:213-230.

Bergman, K., and W. Bonner. 1976. Analysis error as a function of observation density for satellite temperature soundings with spatially correlated errors. *Mon. Wea. Rev.* **104**:1308-1316.

Bernstein, R. L., and D. B. Chelton. 1985. Large-scale sea surface temperature variability from satellite and shipboard measurements. *J. Geophys. Res.* **90**:11619-11630.

Botts, M. E., ed. 1992. *The State of Scientific Visualization with Regard to the NASA EOS Mission to Planet Earth.* Report to NASA Headquarters. In press.

Bower, A. S., and T. Rossby. 1989. Evidence of cross-frontal exchange processes in the Gulf Stream based on isopycnal RAFOS float data. *J. Phys. Oceanogr.* **19**:1177-1190.

Bratseth, A. M. 1986. Statistical interpolation by means of successive corrections. *Tellus, Ser. A* **38**:439-447.

Bretherton, F., R. Davis, and C. Fandry. 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. *Deep-Sea Res.* **23**:559-582.

Brillinger, D. R., and M. Rosenblatt. 1967. Asymptotic theory of kth order spectra. Pp. 153-188 in *Spectral Analysis of Time Series*. B. Harris, ed. Wiley, New York.

Brown, O. B., P. C. Cornillon, S. R. Emerson, and H. M. Carle. 1986. Gulf Stream warm-core rings: a statistical study of their behavior. *Deep-Sea Res.* **33**:1459-1473.

Bryan, K. 1969. A numerical method for the study of the circulation of the World Ocean. *J. Comput. Phys.* **4**:347-376.

Bryson, A. E., and Y. C. Ho. 1975. *Applied Optimal Control.* Hemisphere Publishing, Washington, D.C. 481 pp.

Chelton, D. B. 1983. Effects of sampling errors in statistical estimation. *Deep-Sea Res.* **30**:1083-1103.

Chelton, D. B., and D. B. Enfield. 1986. Ocean signals in tide gauge records. *J. Geophys. Res.* **91**:9081-9098.

Chelton, D. B., and M. G. Schlax. 1991. Estimation of time averages from irregularly spaced observations: with application to Coastal Zone Color Scanner estimates of chlorophyll concentration. *J. Geophys. Res.* **96**:14669-14692.

Collins, M. J., and W. J. Emery. 1988. A computational method for estimating sea ice motion in sequential Seasat synthetic aperture radar imagery by matched filtering. *J. Geophys. Res.* **93**:9251.

Cornillon, P. 1986. The effect of the New England seamounts on Gulf Stream meandering as observed from satellite IR imagery. *J. Phys. Oceanogr.* **16**:386-398.

Cornillon, P., and D. R. Watts. 1987. Satellite thermal infrared imagery and inverted echo sounder determinations of the Gulf Stream northern edge. *J. Atmos. Oceanic Technol.* **4**:712-723.

Cornillon, P., R. Weyer, and G. Flier. 1989. Translational velocity of warm-core rings relative to the slope water. *J. Phys. Oceanogr.* **19**:1317-1332.

Corrsin, S. 1975. Random geometric problems suggested by turbulence. Pp. 300-316 in *Statistical Models and Turbulence.* Lecture Notes in Physics, Vol. 12. M. Rosenblatt and Van Atta, eds. Springer-Verlag, New York.

Cox, M. D. 1984. A primitive equation, three-dimensional model of the ocean. Geophysical Fluid Dynamics Laboratory Ocean Group Tech. Rept. No. 1. GFDL/Princeton University, Princeton, N.J.

Crank, K. N. 1993. NSF program announcement: Collaborative research in geosciences, geography and mathematical sciences. *IMS Bull.* **22**(1):9-10.

Daley, R. 1991. *Atmospheric Data Analysis.* Cambridge University Press, Cambridge.

Davis, R. E. 1976. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.* **6**:249-266.

Davis, R. E. 1977. Techniques for statistical analysis and prediction of geophysical fluid systems. *Geophys. Astrophys. Fluid Dyn.* **8**:245-277.

Davis, R. E. 1983. Oceanic property transport, Lagrangian particle statistics, and their prediction. *J. Mar. Res.* **41**:163-194.

Davis, R. E. 1985. Drifter observations of coastal surface currents during CODE: the statistical and dynamical view. *J. Geophys. Res.* **90**:4756-4772.

Davis, R. E. 1987. Modelling eddy transport of passive tracers. *J. Mar. Res.* **45**:635-666.

Davis, R. E. 1991a. Lagrangian Ocean Studies. *Ann. Rev. Fluid Mech.* **23**:43-64.

Davis, R. E. 1991b. Observing the general circulation with floats. *Deep-Sea Res.* **38**:5531-5571.

Dickey, T. D. 1990. Physical-optical-biological scales relevant to recruitment in large marine ecosystems. Pp. 82-98 in *Large Marine Ecosystems: Patterns, Processes, and Yields.* K. Sherman, L. M. Alexander, and B. D. Gold, eds. AAAS, Washington, D.C.

Dickey, T. D. 1991. The emergence of concurrent high-resolution physical and bio-optical measurements in the upper ocean and their applications. *Rev. Geophys.* **29**(3):383-413.

Douglas, B. C., and R. E. Cheney. 1990. Geosat: Beginning a new era in satellite oceanography. *J. Geophys. Res.* **95**:2833-2836.

Dutkiewicz, S., A. Griffa, and D. B. Olson. 1993. Particle diffusion in a meandering jet. *J. Geophys. Res.* In press.

Dutton, J. A. 1989. The EOS data and information system: Concepts for design. *IEEE Trans. Geosci. Remote Sensing* **27**:109-116.

Emery, W. J., A. C. Thomas, M. J. Collins, W. R. Crawford, and D. L. Mackas. 1986. An objective method for computing advective surface velocities from sequential infrared satellite images. *J. Geophys. Res.* **91**:12865-12878.

Emery, W. J., C. Fowler, and C. A. Clayson. 1992. Satellite-image-derived Gulf Stream currents compared with numerical model results. *J. Atmos. Oceanic Technol.* **9**:286-304.

Farge, M. 1992. Wavelet transforms and their application to turbulence. *Ann. Rev. Fluid Mech.* **24**:395-457.

Figueroa, H. A., and D. B. Olson. 1989. Lagrangian statistics in the South Atlantic as derived from SOS and FGCE drifters. *J. Mar. Res.* **47**:525-546.

Fily, M., and D. A. Rothrock. 1990. Opening and closing of sea ice leads: digital measurements from synthetic aperture radar. *J. Geophys. Res.* **95**:789-796.

Florchinger, P., and F. LeGland. 1990. Time-discretization of the Zakai equation for diffusion processes observed in correlated noise. *Rapp. Rech.-IRIA* **1222**. (Also in *Stochastics and Stochastic Rpts.* **35**(4):233-256, 1991.)

Freilich, M. H., and D. B. Chelton. 1986. Wavenumber spectra of Pacific winds measured by the SEASAT scatterometer. *J. Phys. Oceanogr.* **16**(4):741-757.

Fu, Lee-Lueng. 1983. On the wave number spectrum of oceanic mesoscale variability observed by the SEASAT altimeter. *J. Geophys. Res.* **88**(C7):4331-4341.

Fukumori, I. J. Benveniste, C. Wunsch, and D. B. Haidvogel. 1993. Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother. *J. Phys. Oceanogr.* In press.

Gandin, L. S. 1965. *Objective Analysis of Meteorological Fields*. Israel Program for Scientific Translations, Jerusalem.

Garcia, C. A. E., and I. S. Robinson. 1989. Sea surface velocities in shallow seas extracted from sequential coastal zone color scanner satellite data. *J. Geophys. Res.* **94**:12681-12691.

Gaspar, P., and C. Wunsch. 1989. Estimates from altimeter data of barotropic Rossby waves in the northwestern Atlantic Ocean. *J. Phys. Oceanogr.* **19**:1821-1844.

Gauthier, P. 1992. Chaos and quasi-dimensional data assimilation: a study based on the Lorenz model. *Tellus* **44A**:2-17.

Ghil, M., and P. Malanotte-Rizzoli. 1991. Data assimilation in meteorology and oceanography. Pp. 141-266 in *Advances in Geophysics* **33**. B. Saltzmann, ed. Academic Press, Orlando, Fla.

Gill, A. E. 1982. *Atmosphere-Ocean Dynamics*. Academic Press, New York.

Glazman, R. E. 1986. Statistical characterization of sea surface geometry from a wave slope field discontinuous in the mean square. *J. Geophys. Res.* **91**(C5):6629-6641.

Glazman, R. E. 1991. Statistical problems of wind-generated gravity waves arising in microwave remote sensing of surface winds. *Trans. Geosci. Remote Sensing* **29**(1):135-142.

Glazman, R. E., and P. Weichman. 1989. Statistical geometry of a small surface patch in a developed sea. *J. Geophys. Res.* **94**(C4):4998-5010.

Glazman, R. E., and M. A. Srokosz. 1991. Equilibrium wave spectrum and sea state bias in satellite altimetry. *J. Phys. Oceanogr.* **21**(11):1609-1621.

Gnanadesikan, R. 1990. Looking ahead: Cross-disciplinary opportunities for statistics. *Am. Stat.* **44**(2):121-125.

Goel, P. K., C. Hagwood, J. R. Kettenring, Y. D. Mittal, and S. C. Schwartz. 1990. Report on cross-disciplinary activities. *IMS Bull.* **19**(6), 716-719.

Gordon, H. R., and D. J. Castano. 1987. The coastal zone color scanner atmospheric correction algorithm: Multiple scattering effects. *Appl. Opt.* **26**:2111-2122.

Graber, H. 1993. A "pattern correlation" method for the analysis of directional ocean wave spectra. In preparation.

Haltiner, G. J., and R. T. Williams. 1980. *Numerical Prediction and Dynamic Meteorology.* Wiley, New York. 477 pp.

Harris, W. C. 1993. Letter from MPS: Significant ideas and proposals must break the boundaries of project barriers. *Amstat News* **197**(April):17.

Hasselmann, K. 1962. On the non-linear energy transfer in a gravity-wave spectrum. *J. Fluid Mech.* **12**:481-500.

Hasselmann, K., W. Munk, and G. MacDonald. 1963. Bispectra of Ocean Waves. Pp. 125-139 in *Proceedings of the Symposium on Time Series Analysis.* M. Rosenblatt, ed. Wiley, New York. 497 pp.

Hoadley, A. B., and J. R. Kettenring. 1990. Communications between statisticians and engineers/physical scientists. *Technometrics* **32**(3):243-247. (With commentaries, pp. 249-274.)

Holland, W. R. 1977. Oceanic general circulation models. Pp. 3-45 in *The Sea*, Vol. 6. E. D. Goldberg, I. N. McCave, J. J. O'Brien, and J. H. Steele, eds. Wiley, New York.

Hunt, J. C. R., O. M. Philips, and D. Williams, eds. 1991. *Turbulence and Stochastic Processes: Kolmogorov's Ideas 50 Years On.* Royal Society of London. Cambridge University Press, Cambridge.

Hwang, C. 1991. *Orthogonal Functions over the Oceans and Applications to the Determination of Orbit Error, Geoid and Sea Surface Topography from Satellite Altimetry.* Report 414. Department of Geodetic Science and Survey, Ohio State University, Columbus.

Institute of Mathematical Statistics. 1988. *Cross-Disciplinary Research in the Statistical Sciences.* Report of a Panel of the IMS. I. Olkin and J. Sacks, co-chairs. 62 pp. (Reprinted in *Statistical Science* **5**(February):121-146, 1990.)

Jackson, F. C. 1979. The reflection of impulses from a nonlinear random sea. *J. Geophys. Res.* **84**(C8):4939-4943.

Journel, A. G., and C. Huijbregts. 1978. *Mining Geostatistics.* Academic Press, San Diego, Calif.

Kamachi, M. 1989. Advective surface velocities derived from sequential images for rotational flow field: limitations and applications of maximum cross correlation method with rotational registration. *J. Geophys. Res.* **94**:18227-18233.

Kelly, K. A. 1989. An inverse model for near-surface velocity from infrared images. *J. Phys. Oceanogr.* **19**:1845-1864.

Kelly, K. A., and P. T. Strub. 1992. Comparison of Velocity Estimates from AVHRR in the Coastal Transition Zone. *J. Geophys. Res.* **97**:9653-9668.

Kenny, B. 1982. Beware of spurious self-correlations! *Water Resour. Res.* **18**:1041-1048.

Kouzai, K., and K. Tsuchiya. 1990. Estimation of surface current velocities of the Kuroshio using single NOAA AVHRR image. *J. Remote Sensing Soc. Japan* **10**(1):13-2X (in Japanese).

Kraus, W., and C. W. Boning. 1987. Lagrangian properties of eddy fields in the northern North Atlantic as deduced from satellite-tracked buoys. *J. Mar. Res.* **45**:259-291.

Li, W., and R. N. Meroney. 1985. Re-examination of Eulerian-Lagrangian turbulence relationships. *Atmos. Environ.* **19**:853-855.

Lii, K. S., and M. Rosenblatt. 1990. Asymptotic normality of cumulant spectral estimates. *J. Theor. Prob.* 3:367-385.

Longuet-Higgins, M. S. 1957. The statistical analysis of a random, moving surface. *Philos. Trans. R. Soc. London, Ser. A.* **249**:321-387.

Longuet-Higgins, M. S. 1962. The statistical geometry of random surfaces. Pp. 105-142 in *Hydrodynamic Instability*. American Mathematical Society Proceedings of Symposia in Applied Mathematics, Vol. XIII. American Mathematical Society, Providence, R.I.

Longuet-Higgins, M. S. 1963. The effect of nonlinearities on statistical distributions in the theory of sea waves. *J. Fluid Mech.* **17**:459-480.

Longuet-Higgins, M. S. 1984. Statistical properties of wave groups in a random sea state. *Philos. Trans. R. Soc. London, Ser. A* **312**:219-250.

Lorenc, A. 1986. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **112**, 1177-1194.

Lorenz, E. N. 1956. Empirical orthogonal functions and statistical weather prediction. Sci. Rept. No. 1. Statistical Forecasting Project. Department of Meteorology, Massachusetts Institute of Technology, Cambridge, Mass. 49 pp.

Lovejoy, S., and D. Schertzer. 1986. Scale invariance, symmetries, fractals and stochastic simulation of atmospheric phenomena. *Bull. Am. Meteorol. Soc.* **67**:21-32.

Major, P. 1980. *Multiple Wiener-Ito Integrals.* Lecture Notes in Mathematics, No. 849. Springer-Verlag, New York. 126 pp.

Mandelbrot, B. 1974. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.* **62**:331-350.

Mariano, A. J. 1990. Contour analysis: a new approach for melding geophysical fields. *J. Atmos. Oceanic Technol.* **7**:285-295.

Maury, M. F. 1855. *The Physical Geography of the Sea and Its Meteorology.* 1st edition, Harper and Bros., New York; 8th edition (1861) republished 1963, J. Leighly, ed., Harvard University Press, Cambridge, Mass.

McCormick, B. H., T. A. DeFanti, and M. D. Brown. 1987. Visualization in scientific computing. *Computer Graphics* **21**(6; Nov.) ACM SIGGRAPH, New York.

McIntosh, P. C. 1990. Oceanographic data interpolation: Objective analysis and splines. *J. Geophys. Res.* **95**:13259-13541.

Miller, R. N., and M. A. Cane. 1989. A Kalman filter analysis of sea level height in the tropical Pacific. *J. Phys. Oceanogr.* **19**:773-790.

Miller, R. N., M. Ghil, and F. Gauthiez. 1993. Data assimilation in non-linear dynamical systems. *J. Atmos. Sci.*, in press.

Molchanov, S., and L. Piterbarg. 1992. Heat propagation in random media. *Russ. J. Math. Phys.*, in press.

Monin, A. S., and A. M. Yaglom. 1971. *Statistical Fluid Mechanics, I.* MIT Press, Cambridge, Mass.

Monin, A. S., and A. M. Yaglom. 1975. *Statistical Fluid Mechanics, II.* MIT Press, Cambridge, Mass.

Moore, A. M. 1991. Data assimilation in a quasi-geostrophic open-ocean model of the Gulf Stream region. *J. Phys. Oceanogr.* **21**:398-427.

Morrow, R., J. Church, R. Coleman, D. Chelton, and N. White. 1992. Eddy momentum flux and its contribution to the Southern Ocean momentum balance. *Nature* **357**(11 June):482-484.

National Research Council (NRC). 1984. *Computational Modeling and Mathematics Applied to the Physical Sciences.* Committee on Applications of Mathematics. National Academy Press, Washington, D.C.

National Research Council (NRC). 1987. *Interdisciplinary Research in Mathematics, Science, and Technology Education.* Committee on

Research in Mathematics, Science, and Technology Education. National Academy Press, Washington, D.C.

National Research Council (NRC). 1988. *Oceanography and the Navy: Future Directions.* Navy Review Panel. Ocean Studies Board. National Academy Press, Washington, D.C.

National Research Council (NRC). 1990a. *Interdisciplinary Research: Promoting Collaboration Between the Life Sciences and Medicine and the Physical Sciences and Engineering.* Committee on Promoting Research Collaboration. National Academy Press, Washington, D.C.

National Research Council (NRC). 1990b. *Sea-Level Change.* Geophysics Study Committee. National Academy Press, Washington, D.C.

National Research Council (NRC). 1991a. *Four-Dimensional Model Assimilation of Data: A Strategy for the Earth System Sciences.* Panel on Model-Assimilated Data Sets for Atmospheric and Oceanic Research. Board on Atmospheric Sciences and Climate. National Academy Press, Washington, D.C.

National Research Council (NRC). 1991b. *Spatial Statistics and Digital Image Analysis.* Panel on Spatial Statistics and Image Processing. Board on Mathematical Sciences. National Academy Press, Washington, D.C.

National Research Council (NRC). 1992a. *Combining Information: Statistical Issues and Opportunities for Research.* Committee on Applied and Theoretical Statistics. Board on Mathematical Sciences. National Academy Press, Washington, D.C.

National Research Council (NRC). 1992b. *Oceanography in the Next Decade: Building New Partnerships.* Ocean Studies Board. National Academy Press, Washington, D.C.

National Research Council (NRC). 1992c. *Probability and Algorithms.* Committee on Applied and Theoretical Statistics. Board on Mathematical Sciences. National Academy Press, Washington, D.C.

Ninnis, R. M., W. J. Emery, and M. J. Collins. 1986. Automated extraction of pack ice motion from advanced very high resolution radiometer imagery. *J. Geophys. Res.* **91**:10725-10734.

Novikov, E. A. 1966. Mathematical models of intermittency in turbulent flow. *Dokl. Akad. Nauk SSSR* **168**:1279-1282.

Novikov, E. A., and R. Stewart. 1964. Intermittency of turbulence and spectrum of fluctuations in energy dissipation. *Izv. Akad. Nauk SSSR, Ser. Geofiz.* **3**:408.

O'Brien, J. J., ed. 1986. *Advanced Physical Oceanographic Numerical Modelling.* Proceedings of NATO Advanced Study Institute on Advanced Physical Oceanographic Numerical Modelling. NATO ASI Series C. 186. Kluwer Academic Press, Norwell, Mass. 608 pp.

Pandey, P. C. 1987. *IEEE Trans. Geosci. Remote Sensing* **GE-25**:6670-6676.

Pedlosky, J. 1987. *Geophysical Fluid Dynamics.* Second edition. Springer-Verlag, New York.

Pond, S., and G. L. Pickard. 1983. *Introductory Dynamic Oceanography.* Peragmon Press, Oxford.

Pool, R. 1992. The third branch of science debuts. *Science* **256**(3 April):44-47.

Preisendorfer, R. W. 1988. *Principal Component Analysis in Meteorology and Oceanography.* Elsevier, New York. 425 pp.

Preisendorfer, R. W., and T. P. Barnett. 1983. Numerical model-reality intercomparison tests using small-sample statistics. *J. Atmos. Sci.* **40**:1884-1896.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1986. *Numerical Recipes.* Cambridge University Press, New York. 818 pp.

Price, J. F., and H. T. Rossby. 1982. Observations of a barotropic planetary wave in the western North Atlantic. *J. Mar. Res.* **40**:543-558.

Ridderinkhof, H., and J. T. Zimmerman. 1992. Chaotic stirring in a tidal system. *Science* **258**:1107-1111.

Ropelewski, C. F. 1992. Predicting El Niño events. *Nature* **356**(9 April):476-477.

Rosenblatt, M. 1985. *Stationary Sequences and Random Fields.* Birkhauser, Boston. 258 pp.

Rosenblatt, M. 1991. *Stochastic Curve Estimation.* NSF-CBMS Series in Probability and Statistics, Vol. 3. Institute of Mathematical Statistics, Hayward, Calif. 93 pp.

Rozovskii, B. L. 1990. *Stochastic Evolution Systems. Linear Theory and Applications to Non-Linear Filtering.* Kluwer Academic Publishers, Dordrecht.

Salby, M. L. 1982. Sampling theory for asymptotic satellite observations, Part I: Space-time spectra, resolution and aliasing. *J. Atmos. Sci.* **39**:2577-2600.

Schertzer, D., and S. Lovejoy. 1987. Physically based rain and cloud modeling by anisotropic multiplicative turbulent cascades. *J. Geophys. Res.* **92**(D8):9693-9714.

Schertzer, D., and S. Lovejoy, eds. 1991. *Non-Linear Variability in Geophysics: Scaling and Fractals.* Kluwer, Boston. 318 pp.

Schlax, M. G., and D. B. Chelton. 1992. Frequency domain diagnostics for linear smoothers. *J. Am. Stat. Assoc.*, in press.

Schmitt, F., D. Lavallee, D. Schertzer, and S. Lovejoy. 1992. Empirical determination of universal multifractal exponents in turbulent velocity fields. *Phys. Rev. Lett.* **69**:305-308.

Semtner, A. J., and R. M. Chervin. 1992. Ocean general circulation from a global eddy-resolving model. *J. Geophys. Res.* **97**:5493-5550.

Silverman, B. W. 1985. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *J. R. Stat. Soc. B* **47**:1-52.

Srokosz, M. A. 1986. On the joint distribution of surface elevation and slopes for a nonlinear random sea, with an application to radar altimetry. *J. Geophys. Res.* **91**(C1):995-1006.

Starr, V. P. 1968. *Physics of Negative Viscosity Phenomena.* McGraw-Hill, New York.

Stern, M. 1975. *Ocean Circulation Physics.* Academic Press, New York. 246 pp.

Sverdrup, H. U., M. W. Johnson, and R. H. Fleming. 1942. *The Oceans: Their Physics, Chemistry and General Biology.* Prentice-Hall, Englewood Cliffs, N.J.

Swain, C. J. 1976. A FORTRAN IV program for interpolating irregularly spaced data using the difference equation for minimum curvature. *Comput. & Geosci.* 1:231-240.

Taylor, G. I. 1915. Eddy motions in the atmosphere. *Philos. Trans. R. Soc. London, Ser. A* **215**:1-26.

Thacker, W. C., and R. B. Long. 1988. filtering dynamics to data. *J. Geophys. Res.* **93**:1227-1240.

Thiebaux, H. J., and M. A. Pedder. 1987. *Spatial Objective Analysis, with Applications in Atmospheric Science.* Academic Press, London.

Thiebaux, H. J., and F. Zwiers. 1984. The interpretation and estimation of effective sample size. *J. Clim. Appl. Meteorol.* **23**:800-811.

Thomson, D. J. 1986. A random walk model of dispersion in turbulent flows and its application to dispersion in a valley. *Q. J. R. Meteorol. Soc.* **112**:511-530.

Tokmakian, R. T., P. T. Strub, and J. McClean-Padman. 1990. Evaluation of the maximum cross-correlation method of estimating sea surface velocities from sequential satellite images. *J. Atmos. Oceanic Technol.* **7**:852-865.

Tripoli, G. J., and T. N. Krishnamurti. 1975. Low-level flows over the GATE area during summer 1972. *Mon. Wea. Rev.* **103**:197-216.

Tziperman, E., and W. C. Thacker. 1989. An optimal-control/adjoint-equations approach to studying the oceanic general circulation. *J. Phys. Oceanogr.* **19**:1471-1485.

Wahl, D. D., and J. J. Simpson. 1991. Satellite derived estimates of the normal and tangential components of near-surface flow. *Int. J. Remote Sensing* **12**:2529-2571.

Walsh, J. 1986. An introduction to stochastic partial differential equations. *Springer Lecture Notes in Mathematics* **1180**:265-439.

Warren, B. A., and C. Wunsch, eds. 1981. *Evolution of Physical Oceanography: Scientific Surveys in Honor of Henry Stommel.* MIT Press, Cambridge, Mass.

Wunsch, C. 1978. The general circulation of the North Atlantic west of 50W determined from inverse methods. *Rev. Geophys.* **16**:583-620.

Wunsch, C. 1988. Transient tracers as a problem in control theory. *J. Geophys. Res.* **93**:8099-8110.

Wunsch, C. 1989. Sampling characteristics of satellite orbits. *J. Atmos. Oceanic Technol.* **6**:891-907.

Yaglom, A. M. 1966. The influence of the fluctuation in energy dissipation on the shape of turbulent characteristics in the inertial interval. *Sov. Phys. Dokl. (Engl. Transl.)* **2**:26-30.

Zakharov, V. E. 1984. Kolmogorov spectra in weak turbulence problems. Pp. 4-36 in *Handbook of Plasma Physics*. M. N. Rosenbluth and R. Z. Sagdeev, eds. Elsevier Science Publishers, New York.

Zakharov, V. E., and V. I. Shrira. 1990. On the formation of the directional spectrum of wind waves. *Zh. Eksp. Teor. Fiz.* **98**(6):1941-1958.

Zakharov, V. E., V. S. L'vov, and G. Falkovich. 1992. *Kolmogorov Spectra of Turbulence: Wave Turbulence.* Springer Verlag, New York. 280 pp.